

AC AAR79163;  
 XX  
 DT 22-DEC-1995 (first entry)  
 XX  
 DE Partial sequence of bovine alpha 3 chain of type IV collagen.  
 XX  
 KW Type IV collagen; alpha 3 chain.  
 XX  
 OS Bos taurus.  
 XX  
 PN US5424408-A.  
 XX  
 PD 13-JUN-1995.  
 XX  
 PR 30-NOV-1990; 90US-0621091.  
 XX  
 PR 30-NOV-1990; 90US-0621091.  
 XX  
 (UNIV ) UNIV KANSAS MEDICAL CENT.  
 (UYVA ) UNIV YALE.  
 PI Hudson BG, Morrison KE, Reeders ST;  
 XX  
 DR WPI: 1995-262631/34.  
 XX  
 N-PSDB; AAQ96290.  
 XX  
 cDNAs encoding human or bovine alpha-3 type 4 collagen peptide(s) -  
 PT useful for detection and therapeutic removal of antibodies associated  
 PT with Goodpasture syndrome  
 XX  
 Disclosure; Columns 5-8; 33pp; English.  
 XX  
 Using the PCR with primers derived from each end of the known 27 AA  
 CC residue bovine alpha 3 (IV) collagen protein sequence, a 68 bp  
 CC bovine genomic fragment was amplified. This fragment was then used  
 CC to a bovine lens cDNA library and a 1.5 kb partial cDNA clone was  
 CC obt'd. (clone KMC15). This encodes 238 residues of the triple helical  
 CC collagenous domain and all 233 residues of the C terminal non-  
 CC collagenous domain (NCL) domain of the alpha 3 (IV) chain. An isolated and  
 CC substantially pure nt. having the sequence in AAQ96290 is claimed.  
 XX  
 Sequence 471 AA:  
 SQ

Query Match 93.7%; Score 986; DB 16; Length 471;  
 Best Local Similarity 92.6%; Pred. No. 6.2e-103; Gaps 0;  
 Matches 176; Conservative 10; Mismatches 4; Indels 0;  
 Gaps 0;

1	ORAHGDLGTIGSCLQRFITMPFLCNVNDVCNFRNDYSWJLSTPALMPMMAPITGR	60
:	:     :     :     :     :     :     :     :     :     :     :	
Db	eqahqdgltqrlqftmpfifcnindvanfasrndyswistpamipmdmapitgr	340
Qy	61 ALEPYTSRCTVEGEGAIAGIAVHSQTDIIPCPHGWISLNKGFSFTMFTSGSEGTTQALA	120
:	:     :     :     :     :     :     :     :     :     :     :     :	
Db	341 amepytsrctvegaaiaavhsqtdipcpaghwiwkgfslfmftagssegqala	400
Qy	121 SFGSCLEFRASPFLECHGRGTCNYNSNSFWIASLNPERRKPKIPSIVKAEBLEKII	180
Db	401 spgscllefraspflechgrgtnyynsnsfswasldpkrmfkpkipstvkagelenii	460
Qy	181 SRCQVCMKKR 190	
Db	461 srcqvcmkr 470	

Search completed: March 6, 2002, 06:56:24  
 Job time: 135 sec

Qy 181 SRCQVCMKKRRH 191  
   |||||  
 Db 208 srcqvcmkrrh 218

RESULT 7  
 AAY44171  
 ID AAY44171 standard; Protein; 471 AA.  
 XX  
 AC AAY44171;  
 XX

DT 01-FEB-2000 (first entry)  
 DE Bovine type IV collagen alpha3 chain protein.  
 XX  
 Recombinant; bovine; alpha3 chain; type IV collagen; detection;  
 KW Goodpasture syndrome; antibody; blood; tissue; human; nephrotropism.  
 XX  
 OS Bos taurus.  
 XX  
 PR US5973120-A.  
 XX  
 PD 26-OCT-1999.  
 PF 07-MAR-1995; 95US-0399889.  
 XX  
 PR 30-NOV-1990; 90US-0621091.  
 PA (UYYA ) UNIV YALE.  
 XX  
 (UNIV ) UNIV KANSAS MEDICAL CENT.  
 PI Hudson BG, Reeder ST, Morrison KE;  
 XX  
 DR WPI; 1999-610317/52.  
 DR N-FSDB; AAZ28774.

XX  
 Isolated alpha 3 chain of type IV collagen polypeptide useful for  
 PT diagnosis and treatment of Goodpasture syndrome -  
 XX  
 PS Claim 1; Column 31-34; 27pp; English.

This sequence represents a recombinant bovine alpha3 chain of type IV  
 CC collagen polypeptide. The sequence corresponds to the 238 amino acids  
 CC of the C-terminal end of the triple helical domain and all 233 amino  
 CC acids of the C-terminal non-collagenous domain. Alpha3 chain collagen  
 CC polypeptides are useful for detecting Goodpasture antibodies in blood  
 CC or tissue from a human patient and for treating Goodpasture syndrome,  
 CC especially by neutralising the antibodies in the blood. The polypeptides  
 XX also have a nephrotrophic activity.  
 Sequence 471 AA;

SQ Query Match 93.9%; Score 988; DB 20; Length 471;  
 Best Local Similarity 93.2%; Pred. No. 3.7e-103;  
 Matches 177; Conservative 9; Mismatches 4; Indels 0; Gaps 0;

Qy 1 QRAHGQDGTGTLGSCLQRTTMMPLFCNVNDVNQFASNDYSWKLSTPALMNNMAPITGR 60  
   :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:  
 Db 281 eqahgqdgltglsclqrttmmplfcnindnfafasndyswklstpalmnnmapitgr 340  
 Qy 61 ALEPYISRCTVEGPAIAIVHSQTIDIPCPCHGWISLWKSEFIMTSAGSEGTOALA 120  
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:  
 Db 341 alepyisrcvcegpaiaivhsqttdippcpagwislwkgsfimtsagsegaggala 400  
 Qy 121 SGSCLEFRASPFLECHGRGTCNYNSYSWLASNPERMERKIPSTVAGELEKII 180  
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:  
 Db 401 spgscleefraspfiechgrgtnyyssysfwlasidpkrmrkripstvkagelenii 460  
 Qy 181 SRCQVCMKKR 190  
   |||||||  
 Db 461 srcqvcmkrr 470

RESULT 9  
 AAR79163  
 ID AAR79163 standard; Protein; 471 AA.  
 XX

Db	258	sreqvcmkkrh	268				
KW	type IV collagen alpha chain monomer; human; inhibitor; angiogenesis; tumour growth; integrin receptor; carcinoma; sarcoma; rhabdomyosarcoma; retinoblastoma; Ewing sarcoma; neuroblastoma; osteosarcoma; Leukaemia; diabetic retinopathy; rheumatoid arthritis; neovascularisation; muscular degeneration; corneal graft rejection; vitamin A deficiency; atopic keratitis; mycobacteria infection; chemical burn; sarcoid; Kaposi's sarcoma; sickle cell anaemia; carotid obstructive disease; chronic inflammation; psoriasis; therapy; alpha3(IV)NC1.						
KW	Homo sapiens.						
XX							
PN	WO200059532-A1.						
XX							
PD	12-OCT-2000.						
XX							
PF	31-MAR-2000; 2000WO-US08678.						
XX							
PR	01-APR-1999; 990US-0121391.						
PS	(BIOS-) BIOSTRATUM INC.						
XX							
PI	Brooks P, Hudson B;						
XX							
WPI:	2000-66462/64.						
DR	N-PsDB; AAQ96293.						
XX							
PT	Use of antagonists of specific integrin receptors for inhibiting angiogenesis, tumour growth or metastases, or endothelial cell interactions with the extracellular matrix						
PT							
XX							
PS	Disclosure; Fig 17C; 78pp; English.						
XX							
CC	This sequence is a human type IV collagen alpha chain monomer, designated alpha3(IV)NC1. The invention relates to a method for inhibiting angiogenesis, tumour growth or metastases, or endothelial cell interactions with the extracellular matrix, comprising contacting the cells or tissue with a polypeptide composition containing antagonists of specific integrin receptors. The methods and the antagonists are useful for inhibiting angiogenesis, tumour growth or metastases, or endothelial cell interaction with the extracellular matrix. The antagonists are also useful for treating diseases and conditions with accompanying undesired angiogenesis, e.g. solid and blood-borne tumours (e.g. melanomas, carcinomas, sarcomas, rhabdomyosarcoma, retinoblastoma, Ewing sarcoma, neuroblastoma, osteosarcoma or leukaemia). These are also applicable to treating non-tumorigenic diseases and conditions with accompanying undesired angiogenesis, e.g. diabetic retinopathy, rheumatoid arthritis, retinal neovascularisation, choroidal neovascularisation, muscular degeneration, corneal graft rejection, uveitis, A deficiency, atopic keratitis, Mycobacteria infections, chemical burns, Kaposi's sarcoma, sickle cell anaemia, sarcoid, carotid obstructive disease, post-laser complications, chronic inflammation or psoriasis.						
CC							
CC	Sequence 268 AA;						
SQ							
Query Match	99 %;	Score	1047;	DB	21;	Length	268;
Best Local Similarity	99.5 %;	Pred. No.	3.7e-110;				
Matches	190;	Conservative	0;	Mismatches	1;	Indels	0;
Gaps							
OY	1	QRAHQDQDGTLGSGCLQRFTTMPELFCNVNDVCFNFAASRNDSYWLSTPALMMNNMAPITGR	60				
Db	78	qrahqdgltlgsgclqrfttmpefncnvdvcafasndyswistpalmmnnmapitgr	137				
OY	61	ALEPYISRCVCEGPAAIAAVHSQTDTIPCPHGWISWKGSFRTMTSAGSEGCGQALA	120				
Db	138	alepyisrtctvcegpaaiaavhsqttdipcpophyislwkgfslmftsdagqala	197				
OY	121	SPGSCLEERASPALECHRGTCWYNSYSFWLASLNPERMFRRKPIPSVKAEGEKKI	180				
Db	198	spgscleerfaspalechrgtcwyyssysfwlaslnpermfrkpiptvakelekkii	257				
OY	181	SRQQVCMKKRH	191				
Query Match	99 %;	Score	1047;	DB	21;	Length	268;
Best Local Similarity	99.5 %;	Pred. No.	3.7e-110;				
Matches	190;	Conservative	0;	Mismatches	1;	Indels	0;
Gaps							
XX							
SQ							
Sequence	218 AA;						
XX							
Query Match	97.4 %;	Score	1025;	DB	16;	Length	218;
Best Local Similarity	97.9 %;	Pred. No.	8.6e-108;				
Matches	187;	Conservative	0;	Mismatches	4;	Indels	0;
Gaps							
OY	1	QRAHQDQDGTLGSGCLQRFTTMPELFCNVNDVCFNFAASRNDSYWLSTPALMMNNMAPITGR	60				
Db	28	qrahqdgltlgsgclqrfttmpefncnvdvcafasndyswistpalmmnnmapitgr	87				
OY	61	ALEPYISRCVCEGPAAIAAVHSQTDTIPCPHGWISWKGSFRTMTSAGSEGCGQALA	120				
Db	88	alepyisrtctvcegpaaiaavhsqttdipcpophyislwkgfslmftsdagqala	147				
OY	121	SPGSCLEERASPALECHRGTCWYNSYSFWLASLNPERMFRRKPIPSVKAEGEKKI	180				

RESULT	4	CC	This sequence represents a recombinant type IV collagen
AAV31993	XX	CC	non-collagenous (NC1) domain alpha-3 polypeptide composed of a
ID	XX	CC	BM40 signal sequence (which is cleaved from the mature protein) to
XX	XX	CC	facilitate protein secretion, and a mature protein comprising an
AC	XX	CC	affinity tag (facilitates purification and identification of the
XX	XX	CC	material) and the alpha-1 chain monomer. The invention provides
XX	XX	CC	methods and kits for inhibiting angiogenesis, tumour growth and
DT	XX	CC	metastasis, and endothelial cell interaction with the extracellular
DE	XX	CC	tissue with 1 or more isolated type IV collagen NC1 alpha chain
XX	XX	CC	monomer(s) selected from the group consisting of alpha-1, alpha-2,
XX	XX	CC	alpha-3 and alpha-6 NC1 chain monomers (see AAV31991-96). The
XX	XX	CC	monomers can be produced via recombinant protein expression. The
XX	XX	CC	polynucleotides and polypeptides are used to treat an angiogenesis-
XX	XX	CC	mediated disorder or condition, especially selected from solid and
XX	XX	CC	blood-borne tumours, diabetic retinopathy, rheumatoid arthritis,
XX	XX	CC	retinal neovascularization, choroidal neovascularization, macular
XX	XX	CC	degeneration, corneal neovascularization, retinopathy of prematurity,
XX	XX	CC	corneal graft rejection, neovascular glaucoma, retrobulbar
XX	XX	CC	fibroplasia, epidemic keratoconjunctivitis, vitamin A deficiency,
XX	XX	CC	contact lens wear, atopic keratitis, superior limbic keratitis,
XX	XX	CC	superior limbic keratitis, sicca, sogrens, acne rosacea, phlyctenulosis,
XX	XX	CC	syphilis, mycobacteria infections, lipid degeneration, chemical
XX	XX	CC	burns, bacterial ulcers, fungal ulcers, herpes simplex infections,
XX	XX	CC	herpes zoster infections, protozoan infections, kaposi's sarcoma,
XX	XX	CC	Mooran ulcer, Terrien's marginal degeneration, marginal keratolysis,
XX	XX	CC	contact lens wear, atopic keratitis, superior limbic keratitis,
XX	XX	CC	scleritis, Steven's Johnson disease, radial keratotomy, sickle cell
XX	XX	CC	anaemia, sarcoid, pseudoxanthoma elasticum, Pagets disease, vein
XX	XX	CC	occlusion, artery occlusion, carotid obstructive disease, chronic
XX	XX	CC	uveitis, chronic vitritis, Lyme's disease, Eales disease, Bechets
XX	XX	CC	disease, myopia, optic pits, Stargardt's disease, pars planitis,
XX	XX	CC	chronic retinal detachment, hyperviscosity syndromes, toxoplasmosis,
XX	XX	CC	post-laser complications, abnormal proliferation of fibrovascular
XX	XX	CC	tissue, haemangioma, Osler-Weber-Rendu, AIDS, ocular neovascular
XX	XX	CC	disease, osteoarthritis, chronic inflammation, Crohn's disease,
XX	XX	CC	ulcerative colitis, psoriasis, atherosclerosis, and pemphigoid (all
OS	SQ	XX	claimed).
OS	SQ	Sequence	268 AA:
XX	XX	FH	Location/Qualifiers
FT	peptide	1..17	/note= "BM40 signal peptide"
FT	Protein	18..268	/note= "mature protein"
FT	peptide	18..25	/note= "affinity tag"
FT	Protein	26..268	/note= "NC1 alpha-3 monomer"
XX	W09949885-A2.	QY	1 QRAHGQDGTGSLQRRTTMLPLFCNVNNDVNFASNDYSYWLSTPALMPNMAPITGR
XX	07-OCT-1999.	Db	78 qrahggdgtlgslqlqftttmbflfcnvnvdvfasrndyswlistpalmpnnmapitgr
XX	26-MAR-1999;	QY	61 ALEPYIISRCTVCEGPAIAVHSQTTDIPPCPCHGWISIWKGSFIMTSAGSEGTGOLA
XX	99WO-0079783.	Db	138 alepyisrltvcgepaiaivhsqtdlippccphgwlslwkgstfimtsasgegala
PR	27-MAR-1998;	QY	121 SPSCLEFRASPFLEGIGRGNYYNSYSFLASINPERMFRKPKPSTVAGELEKII
PR	29-OCT-1998;	Db	198 spgscleefraspflechrgtgcnyyssysfwlasinpermfrkpkpstvkagelekkii
XX	(UNIV ) UNIV KANSAS MEDICAL CENT.	QY	257
PA	Hudson BG, Sarras MP;	Db	181 SRQCVCMKKRH
XX	WPI; 1999-601297/51.	DB	268 srqcvcmkkrh
DR	N-PSDB; AA20091.	DB	268
PT	Inhibition of angiogenesis with non-collagenous alpha chain monomer	XX	
PT	useful for treating e.g. tumor growth or metastasis,	XX	
PT	neovascularisation, etc	XX	
PS	Disclosure; Fig 17c; 56pp; English.	XX	
XX	RESULT	5	
XX	ID	AAV97555	
XX	AAV97555 standard; Protein; 268 AA.	XX	
XX	AC	AAV97555;	
XX	DT	12-FEB-2001 (first entry)	
XX	DE	Human alpha3(IV)NC1 protein sequence.	

Db 115 alepyisrctvcegpaiaivavnsqtdipcpghwislwgfisimtsaqsegaggala 174  
 Qy 121 SPGSCLEEFRAFPFLECHGRGTCNYNSYSWLASINPERMFRKIPSTVKAGELEKII 180  
 Db 175 spgscleefraspflechgrtgctnyysnsyfwtaslnpermfrkipstvkageleki 234  
 Qy 181 SRCQVCMKKRH 191  
 Db 235 srcqvcmkkrh 245

RESULT 2  
 AAY4172  
 ID AAY4172 standard; Protein; 218 AA.  
 AC AAY4172;  
 DT 01-FEB-2000 (first entry)  
 Human type IV collagen alpha3 chain protein.

XX Recombinant; bovine; alpha3 chain; type IV collagen; detection;  
 KW Goodpasture syndrome; antibody; blood; tissue; human; nephrotrophism.  
 XX Homo sapiens.  
 OS US5973120-A.  
 XX PD 26-OCT-1999.  
 PF 07-MAR-1995; 95US-0399889.  
 XX PR 30-NOV-1990; 90US-0621091.  
 XX PA (UYA ) UNIV YALE.  
 PA (UNIV ) UNIV KANSAS MEDICAL CENT.  
 XX PI hudson BG, Reeder ST, Morrison KE.  
 DR WPI; 1999-610317/52.  
 DR N-PSDB; AAZ28775.

XX Isolated alpha 3 chain of type IV collagen polypeptide useful for  
 PT diagnosis and treatment of Goodpasture syndrome -  
 PS Claim 2; Column 35-36; 27pp; English.

CC this sequence represents a recombinant human alpha3 chain of type IV  
 CC collagen polypeptide. The sequence corresponds to the 218 amino acids  
 CC of the C-terminal non-collagenous domain. Alpha3 chain collagen  
 CC polypeptides are useful for detecting Goodpasture antibodies in blood  
 CC or tissue from a human patient and for treating Goodpasture syndrome,  
 CC especially by neutralising the antibodies in the blood. The polypeptides  
 CC also have a nephrotrophic activity.

XX Sequence 218 AA;

SQ Query Match 99.5%; Score 1047; DB 20; Length 218;  
 Best Local Similarity 99.5%; Pred. No. 2.8e-110; Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 ORAHGQDGLTGLGSCLQRTTMPLFCNNNDVNFASRNDYSWLSTPALMPNMAPTR 60  
 Qy 1 ORAHGQDGLTGLGSCLQRTTMPLFCNNNDVNFASRNDYSWLSTPALMPNMAPTR 60  
 Db 28 qrahggdgltgslcqlfttmpflfcvndvenfasndyswlstpalmppnmaptr 87  
 Qy 61 ALEPYISRCTVCEGPAIAIVAVNSQTDIPCPGHWISLWGFSFIMTSAGSEGTOALA 120  
 Db 88 alepyisrctvcegpaiaivavnsqtdipcpghwislwgfisimtsaqsegaggala 147  
 Qy 121 SPGSCLEEFRAFPFLECHGRGTCNYNSYSWLASINPERMFRKIPSTVKAGELEKII 180  
 Db 148 spgscleefraspflechgrtgctnyysnsyfwtaslnpermfrkipstvkageleki 207  
 Qy 181 SRCQVCMKKRH 191  
 Db 208 srcqvcmkkrh 218

GenCore version 4.5  
copyright (c) 1993 - 2000 Compugen Ltd.

## ALIGNMENTS

OM protein - protein search, using sw model  
Run on: March 6, 2002, 06:54:09 ; Search time 54.67 Seconds  
(without alignments)  
Scoring table: BLOSUM62  
Gapext 10.0 , Gapext 0.5  
Searched: 522463 seqs, 74073290 residues  
To number of hits satisfying chosen parameters: 9  
Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 90%  
Maximum Match 100%  
Listing first 1000 summaries

Database :  
1: /SIDS8/geodata/geneseq/geneseq/AA1980.DAT:  
2: /SIDS8/geodata/geneseq/geneseq/AA1981.DAT:  
3: /SIDS8/geodata/geneseq/geneseq/AA1982.DAT:  
4: /SIDS8/geodata/geneseq/geneseq/AA1983.DAT:  
5: /SIDS8/geodata/geneseq/geneseq/AA1984.DAT:  
6: /SIDS8/geodata/geneseq/geneseq/AA1985.DAT:  
7: /SIDS8/geodata/geneseq/geneseq/AA1986.DAT:  
8: /SIDS8/geodata/geneseq/geneseq/AA1987.DAT:  
9: /SIDS8/geodata/geneseq/geneseq/AA1988.DAT:  
10: /SIDS8/geodata/geneseq/geneseq/AA1989.DAT:  
11: /SIDS8/geodata/geneseq/geneseq/AA1990.DAT:  
12: /SIDS8/geodata/geneseq/geneseq/AA1991.DAT:  
13: /SIDS8/geodata/geneseq/geneseq/AA1992.DAT:  
14: /SIDS8/geodata/geneseq/geneseq/AA1993.DAT:  
15: /SIDS8/geodata/geneseq/geneseq/AA1994.DAT:  
16: /SIDS8/geodata/geneseq/geneseq/AA1995.DAT:  
17: /SIDS8/geodata/geneseq/geneseq/AA1996.DAT:  
18: /SIDS8/geodata/geneseq/geneseq/AA1997.DAT:  
19: /SIDS8/geodata/geneseq/geneseq/AA1998.DAT:  
20: /SIDS8/geodata/geneseq/geneseq/AA1999.DAT:  
21: /SIDS8/geodata/geneseq/geneseq/AA2000.DAT:  
22: /SIDS8/geodata/geneseq/geneseq/AA2001.DAT:  
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1052	100.0	245	Human type IV collagen
2	1047	99.5	218	Human type IV collagen
3	1047	99.5	218	Human alpha3 type IV collagen N
4	1047	99.5	268	Human alpha3 (IV)NC
5	1047	99.5	268	partial sequence o
6	1025	97.4	218	Bovine type IV col
7	988	93.9	471	Bovine alpha3 type
8	988	93.9	471	Partial sequence o
9	986	93.7	471	AAR7164

RESULT 1  
AY67942 standard; protein; 245 AA.  
ID AY67942;  
XX  
XX  
XX  
XX  
DE Human type IV collagen alpha 3 chain protein sequence SEQ ID NO:10.  
XX Human; type IV collagen; anti-angiogenic; angiogenesis; cancer;  
KW benign tumour; rheumatoid arthritis; diabetic retinopathy; psoriasis;  
KW ocular angiogenesis disease; Osler-Webber Syndrome; telangiectasia;  
KW myocardial angiogenesis; Plaque neovascularisation; angiofibroma;  
KW atherosclerosis; sclerodema; hypertrophic scar; cat scratch disease;  
KW contraception; obesity.  
XX Homo sapiens.  
XX WO965940-A1.  
XX PR 23-DEC-1999.  
XX PR 17-JUN-1999; 99WO-US13737.  
XX PR 17-JUN-1998; 98US-0089689.  
XX PR 25-MAR-1999; 99US-0126175.  
XX DR XX  
PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.  
XX PT Kalluri R;  
XX PR WPI; 2000-097708/08.  
XX PS N-PSDB; A257158.  
XX  
CC The present sequence represents the human type IV collagen alpha 1, 2 or  
3 chain of Type IV collagen used in, e.g. treatment of benign tumors  
and rheumatoid arthritis.  
XX  
CC Claim 32; Fig 16B; 117pp; English.  
CC The present invention describes an isolated protein chosen from the NCL  
domain of the alpha 1, alpha 2 or alpha 3 chains of type IV collagen or  
a fragment, analogue, derivative or mutant, which has anti-angiogenic  
properties. The anti-angiogenic proteins, multimers and chimeras are  
useful for inhibiting angiogenic activity in mammalian tissue,  
especially for treating diseases chosen from angiogenesis dependent  
cancers, benign tumours, rheumatoid arthritis, diabetic retinopathy,  
psoriasis, ocular angiogenesis, plaque neovascularisation, telangiectasia,  
myocardial angiogenesis, haemophilic joints, angiofibroma, wound granulation, intestinal  
adhesions, atherosclerosis, sclerodema, hypertrophic scars, cat scratch  
disease, Helicobacter pylori ulcers, dialysis graft vascular access  
stenosis, contraception and obesity. The compositions can be used to  
inhibit a disease characterised by angiogenic activity, in conjunction  
with radiation therapy, chemotherapy or immunotherapy.  
CC  
XX Sequence 245 AA;

Query Match 100.0%; Score 1052; DB 21; Length 245;  
Best Local Similarity 100.0%; pred. No. 8.8e-111; Indels 0; Gaps 0;  
Matches 191; Conservative 0; Mismatches 0; Gaps 0;

QY 1 ORAHDGDLGLGSCLQRFTTMLPLFCNVNDVCFNFSRNDYSYWLSTPALMPNNMAPITGR 60  
Db 55 qrahggdgltgsgclqrfttmlplfcnvndvfnasrndysywlstpalmpnnmapitgr 114  
QY 61 ALEPYIISRCTVCERGPATAIAVHSQTTDIPPCPHGWISLWKFSPFMTSASSEGTCGALA 120

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

Om protein - protein search, using sw model  
Run on: March 6, 2002, 06:54:54 ; Search time 50.56 seconds  
(without alignments)  
552.571 Million cell updates/sec

Title: US-09-543-371-10\_COPY\_54\_244  
Perfect score: 1052  
Sequence: 1 QRAHQDQDITGLGSCLQRFTT.....KAGELEKTIISRQCOVCMKKH 191

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 14627329 residues.

T<sub>c</sub> number of hits satisfying chosen parameters: 8

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000  
Post-processing: Minimum Match 90%  
Maximum Match 100%

Listing first 1000 summaries  
SPTREMBL17:  
1: sp\_archaea:  
2: sp\_bacteria:  
3: sp\_fungi:  
4: sp\_human:  
5: sp\_invertebrate:  
6: sp\_invertebrate:  
7: sp\_mhc:  
8: sp\_organelle:  
9: sp\_phage:  
10: sp\_plant:  
11: sp\_rabbit:  
12: sp\_virus:  
13: sp\_vertebrate:  
14: sp\_unclassified:  
\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	1052	100.0	245	4	Q9NYC4	Q9nyc4 homo sapien	
2	1052	100.0	1670	4	Q9BQT2	Q9bqt2 homo sapien	
3	1043	99.1	212	6	Q28512	Q28512 macaca mulatta	
4	1000	95.1	212	6	Q28567	Q28567 ovis aries	
5	996	94.7	246	11	Q61435	[1] Q61435 mus musculus	
6	996	94.7	1669	11	Q9Q250	Q9q250 mus musculus	
7	990	94.1	230	11	Q63122	Q63122 rattus norvegicus	
8	963	91.5	203	6	Q28682	Q28682 oryctolagus cuniculus	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	1052	100.0	245	4	Q9NYC4	Q9nyc4 homo sapien	
2	1052	100.0	1670	4	Q9BQT2	Q9bqt2 homo sapien	
3	1043	99.1	212	6	Q28512	Q28512 macaca mulatta	
4	1000	95.1	212	6	Q28567	Q28567 ovis aries	
5	996	94.7	246	11	Q61435	[1] Q61435 mus musculus	
6	996	94.7	1669	11	Q9Q250	Q9q250 mus musculus	
7	990	94.1	230	11	Q63122	Q63122 rattus norvegicus	
8	963	91.5	203	6	Q28682	Q28682 oryctolagus cuniculus	

## ALIGNMENTS

RESULT	1	Q9NYC4	PRELIMINARY;	PRT;	245 AA.
ID	Q9NYC4				
AC					
DT	01-OCT-2000	(TREMBLrel. 15, Created)			
DT	01-OCT-2000	(TREMBLrel. 15, Last sequence update)			
RESULT	2	Q9BQT2	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670 AA.
ID	Q9BQT2;				
AC					
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
RESULT	2	Q9BQT2;	PRELIMINARY;	PRT;	1670

DR EMBL: AJ288494; CAC36101..1; JOINED.  
 DR EMBL: AJ288495; CAC36101..1; JOINED.  
 DR EMBL: AJ288496; CAC36101..1; JOINED.  
 DR EMBL: AJ288497; CAC36101..1; JOINED.  
 DR EMBL: AJ288498; CAC36101..1; JOINED.  
 DR EMBL: AJ288499; CAC36101..1; JOINED.  
 DR EMBL: AJ288500; CAC36101..1; JOINED.  
 DR EMBL: AJ288501; CAC36101..1; JOINED.  
 DR EMBL: AJ288502; CAC36101..1; JOINED.  
 DR EMBL: AJ288503; CAC36101..1; JOINED.  
 DR EMBL: AJ288504; CAC36101..1; JOINED.  
 DR EMBL: AJ288505; CAC36101..1; JOINED.  
 DR EMBL: AJ288506; CAC36101..1; JOINED.  
 DR EMBL: AJ288507; CAC36101..1; JOINED.  
 DR EMBL: AJ288508; CAC36101..1; JOINED.  
 DR EMBL: AJ288509; CAC36101..1; JOINED.  
 DR EMBL: AJ288510; CAC36101..1; JOINED.  
 DR EMBL: AJ288511; CAC36101..1; JOINED.  
 DR EMBL: AJ288512; CAC36101..1; JOINED.  
 DR EMBL: AJ288513; CAC36101..1; JOINED.  
 DR EMBL: AJ288514; CAC36101..1; JOINED.  
 DR EMBL: AJ288515; CAC36101..1; JOINED.  
 DR EMBL: AJ288516; CAC36101..1; JOINED.  
 DR EMBL: AJ288517; CAC36101..1; JOINED.  
 DR EMBL: AJ288518; CAC36101..1; JOINED.  
 DR EMBL: AJ288519; CAC36101..1; JOINED.  
 DR EMBL: AJ288520; CAC36101..1; JOINED.  
 DR EMBL: AJ288521; CAC36101..1; JOINED.  
 DR EMBL: AJ288522; CAC36101..1; JOINED.  
 DR EMBL: AJ288523; CAC36101..1; JOINED.  
 DR EMBL: AJ288524; CAC36101..1; JOINED.  
 DR EMBL: AJ288525; CAC36101..1; JOINED.  
 DR EMBL: AJ288526; CAC36101..1; JOINED.  
 DR EMBL: AJ288527; CAC36101..1; JOINED.  
 DR EMBL: AJ288528; CAC36101..1; JOINED.  
 DR EMBL: AJ288529; CAC36101..1; JOINED.  
 DR EMBL: AJ288530; CAC36101..1; JOINED.  
 DR EMBL: AJ288531; CAC36101..1; JOINED.  
 DR EMBL: AJ288532; CAC36101..1; JOINED.  
 DR EMBL: AJ288533; CAC36101..1; JOINED.  
 DR EMBL: AJ288534; CAC36101..1; JOINED.  
 DR EMBL: AJ288535; CAC36101..1; JOINED.  
 DR EMBL: AJ288536; CAC36101..1; JOINED.  
 DR EMBL: AJ288537; CAC36101..1; JOINED.  
 DR EMBL: AJ288538; CAC36101..1; JOINED.  
 KW Collagen. SQ 1670 AA: 161899 MW: FA7BE4914CA0A6F6 CRC64;

Query Match Score 100.0%; Pred. No. 1.e-100; Length 1670;  
 Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;

RESULT 4  
 Q28567 ID Q28567 PRELIMINARY; PRT; 212 AA.  
 AC Q28567 ID Q28567 PRELIMINARY; PRT; 212 AA.

DR 1 QRAHGDGLTGSLCLOFTMPFLCNVNDVCNFASRNDYSYWLSTPALMNNAPITGR 60  
 DR 1480 QRAHGDGLTGSLCLOFTMPFLCNVNDVCNFASRNDYSYWLSTPALMNNAPITGR 1539  
 DR 61 ALEPYISRCTVCEPGLATAIVHSQTFDIPPOPHGNLWKGFSFIMFTSAGSEGIGQALA 120  
 DR 82 ALEPYISRCTVCEPGLATAIVHSQTFDIPPOPHGNLWKGFSFIMFTSAGSEGIGQALA 141  
 DR 121 SPGSCLEEFRASPFLCHGRCTCNYSNSYSFWLASLNPERMFRKPIPSVYKAGELEKII 180  
 DR 142 SPGSCLEEFRASPFLCHGRCTCNYSNSYSFWLASLNPERMFRKPIPSVYKAGELEKII 201  
 QY 181 SRCQVCMKKRH 191  
 DB 202 SRCQVCMKKRH 212

DR EMBL: AJ288539; CAC36101..1; JOINED.  
 DR EMBL: AJ288540; CAC36101..1; JOINED.  
 DR EMBL: AJ288541; CAC36101..1; JOINED.  
 DR EMBL: AJ288542; CAC36101..1; JOINED.  
 DR EMBL: AJ288543; CAC36101..1; JOINED.  
 DR EMBL: AJ288544; CAC36101..1; JOINED.  
 DR EMBL: AJ288545; CAC36101..1; JOINED.  
 DR EMBL: AJ288546; CAC36101..1; JOINED.  
 DR EMBL: AJ288547; CAC36101..1; JOINED.  
 DR EMBL: AJ288548; CAC36101..1; JOINED.  
 DR EMBL: AJ288549; CAC36101..1; JOINED.  
 DR EMBL: AJ288550; CAC36101..1; JOINED.  
 DR EMBL: AJ288551; CAC36101..1; JOINED.  
 DR EMBL: AJ288552; CAC36101..1; JOINED.  
 DR EMBL: AJ288553; CAC36101..1; JOINED.  
 DR EMBL: AJ288554; CAC36101..1; JOINED.  
 DR EMBL: AJ288555; CAC36101..1; JOINED.  
 DR EMBL: AJ288556; CAC36101..1; JOINED.  
 DR EMBL: AJ288557; CAC36101..1; JOINED.  
 DR EMBL: AJ288558; CAC36101..1; JOINED.  
 DR EMBL: AJ288559; CAC36101..1; JOINED.  
 DR EMBL: AJ288560; CAC36101..1; JOINED.  
 DR EMBL: AJ288561; CAC36101..1; JOINED.  
 DR EMBL: AJ288562; CAC36101..1; JOINED.  
 DR EMBL: AJ288563; CAC36101..1; JOINED.  
 DR EMBL: AJ288564; CAC36101..1; JOINED.  
 DR EMBL: AJ288565; CAC36101..1; JOINED.  
 DR EMBL: AJ288566; CAC36101..1; JOINED.  
 DR EMBL: AJ288567; CAC36101..1; JOINED.  
 DR EMBL: AJ288568; CAC36101..1; JOINED.  
 DR EMBL: AJ288569; CAC36101..1; JOINED.  
 DR EMBL: AJ288570; CAC36101..1; JOINED.  
 DR EMBL: AJ288571; CAC36101..1; JOINED.  
 DR EMBL: AJ288572; CAC36101..1; JOINED.  
 DR EMBL: AJ288573; CAC36101..1; JOINED.  
 DR EMBL: AJ288574; CAC36101..1; JOINED.  
 DR EMBL: AJ288575; CAC36101..1; JOINED.  
 DR EMBL: AJ288576; CAC36101..1; JOINED.  
 DR EMBL: AJ288577; CAC36101..1; JOINED.  
 DR EMBL: AJ288578; CAC36101..1; JOINED.  
 DR EMBL: AJ288579; CAC36101..1; JOINED.  
 DR EMBL: AJ288580; CAC36101..1; JOINED.  
 DR EMBL: AJ288581; CAC36101..1; JOINED.  
 DR EMBL: AJ288582; CAC36101..1; JOINED.  
 DR EMBL: AJ288583; CAC36101..1; JOINED.  
 DR EMBL: AJ288584; CAC36101..1; JOINED.  
 DR EMBL: AJ288585; CAC36101..1; JOINED.  
 DR EMBL: AJ288586; CAC36101..1; JOINED.  
 DR EMBL: AJ288587; CAC36101..1; JOINED.  
 DR EMBL: AJ288588; CAC36101..1; JOINED.  
 DR EMBL: AJ288589; CAC36101..1; JOINED.  
 DR EMBL: AJ288590; CAC36101..1; JOINED.  
 DR EMBL: AJ288591; CAC36101..1; JOINED.  
 DR EMBL: AJ288592; CAC36101..1; JOINED.  
 DR EMBL: AJ288593; CAC36101..1; JOINED.  
 DR EMBL: AJ288594; CAC36101..1; JOINED.  
 DR EMBL: AJ288595; CAC36101..1; JOINED.  
 DR EMBL: AJ288596; CAC36101..1; JOINED.  
 DR EMBL: AJ288597; CAC36101..1; JOINED.  
 DR EMBL: AJ288598; CAC36101..1; JOINED.  
 DR EMBL: AJ288599; CAC36101..1; JOINED.  
 DR EMBL: AJ288600; CAC36101..1; JOINED.  
 DR EMBL: AJ288601; CAC36101..1; JOINED.  
 DR EMBL: AJ288602; CAC36101..1; JOINED.  
 DR EMBL: AJ288603; CAC36101..1; JOINED.  
 DR EMBL: AJ288604; CAC36101..1; JOINED.  
 DR EMBL: AJ288605; CAC36101..1; JOINED.  
 DR EMBL: AJ288606; CAC36101..1; JOINED.  
 DR EMBL: AJ288607; CAC36101..1; JOINED.  
 DR EMBL: AJ288608; CAC36101..1; JOINED.  
 DR EMBL: AJ288609; CAC36101..1; JOINED.  
 DR EMBL: AJ288610; CAC36101..1; JOINED.  
 DR EMBL: AJ288611; CAC36101..1; JOINED.  
 DR EMBL: AJ288612; CAC36101..1; JOINED.  
 DR EMBL: AJ288613; CAC36101..1; JOINED.  
 DR EMBL: AJ288614; CAC36101..1; JOINED.  
 DR EMBL: AJ288615; CAC36101..1; JOINED.  
 DR EMBL: AJ288616; CAC36101..1; JOINED.  
 DR EMBL: AJ288617; CAC36101..1; JOINED.  
 DR EMBL: AJ288618; CAC36101..1; JOINED.  
 DR EMBL: AJ288619; CAC36101..1; JOINED.  
 DR EMBL: AJ288620; CAC36101..1; JOINED.  
 DR EMBL: AJ288621; CAC36101..1; JOINED.  
 DR EMBL: AJ288622; CAC36101..1; JOINED.  
 DR EMBL: AJ288623; CAC36101..1; JOINED.  
 DR EMBL: AJ288624; CAC36101..1; JOINED.  
 DR EMBL: AJ288625; CAC36101..1; JOINED.  
 DR EMBL: AJ288626; CAC36101..1; JOINED.  
 DR EMBL: AJ288627; CAC36101..1; JOINED.  
 DR EMBL: AJ288628; CAC36101..1; JOINED.  
 DR EMBL: AJ288629; CAC36101..1; JOINED.  
 DR EMBL: AJ288630; CAC36101..1; JOINED.  
 DR EMBL: AJ288631; CAC36101..1; JOINED.  
 DR EMBL: AJ288632; CAC36101..1; JOINED.  
 DR EMBL: AJ288633; CAC36101..1; JOINED.  
 DR EMBL: AJ288634; CAC36101..1; JOINED.  
 DR EMBL: AJ288635; CAC36101..1; JOINED.  
 DR EMBL: AJ288636; CAC36101..1; JOINED.  
 DR EMBL: AJ288637; CAC36101..1; JOINED.  
 DR EMBL: AJ288638; CAC36101..1; JOINED.  
 KW Collagen. SQ 1670 AA: 161899 MW: FA7BE4914CA0A6F6 CRC64;

Query Match Score 100.0%; Pred. No. 1.e-100; Length 1670;  
 Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;

RESULT 4  
 Q28567 ID Q28567 PRELIMINARY; PRT; 212 AA.

DR 1 QRAHGDGLTGSLCLOFTMPFLCNVNDVCNFASRNDYSYWLSTPALMNNAPITGR 60  
 DR 1480 QRAHGDGLTGSLCLOFTMPFLCNVNDVCNFASRNDYSYWLSTPALMNNAPITGR 1539  
 DR 61 ALEPYISRCTVCEPGLATAIVHSQTFDIPPOPHGNLWKGFSFIMFTSAGSEGIGQALA 120  
 DR 1540 ALEPYISRCTVCEPGLATAIVHSQTFDIPPOPHGNLWKGFSFIMFTSAGSEGIGQALA 1599  
 DR 121 SPGSCLEEFRASPFLCHGRCTCNYSNSYSFWLASLNPERMFRKPIPSVYKAGELEKII 180  
 DR 1600 SGSCLEEFRASPFLCHGRCTCNYSNSYSFWLASLNPERMFRKPIPSVYKAGELEKII 1659  
 DR 181 SRCQVCMKKRH 191  
 DR 1660 SRCQVCMKKRH 1670  
 PRT; 212 AA.

RESULT 3  
 Q28512 ID Q28512 PRELIMINARY; PRT; 212 AA.

DR 01-NOV-1996 (TREMBLrel. 01, Created)  
 DR 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE ALPHA-3 TYPE IV COLLAGEN (FRAGMENT).  
 GN COL4A3.  
 OS Macaca mulatta (Rhesus macaque).  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Macaca.  
 NCBI\_TAXID=9544;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-RENAL CORTEX;  
 RA Turner A.N., Ryan J.J., Derry C.J., Cashman S.J., Kattabma I.,  
 RA Mason P.J., Pusey C.D.;  
 RA Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: L47280; AAA91861..1; -.  
 DR InterPro; IPR00504; RRM.  
 DR InterPro; IPR001442; C4.  
 DR Pfam; PF01413..C4; 2.  
 DR ProdDom; PD003923; C4; 2.  
 DR PROSITE; PS00030; RNP\_1; UNKNOWN\_1.  
 DR SMART; SM0011..C4; 2.  
 FT NON-TER 1  
 FT NON-TER 1  
 SQ SEQUENCE 212 AA; 23469 MW; 4BC574A64E357B64 CRC64;

Query Match Score 99.1%; Pred. No. 1.e-100; Length 212;  
 Best Local Similarity 99.0%; Mismatches 1; Indels 0; Gaps 0;

DR 1 QRAHGDGLTGSLCLOFTMPFLCNVNDVCNFASRNDYSYWLSTPALMNNAPITGR 60  
 DR 22 ERAHGDDLTGSLCLOFTMPFLCNVNDVCNFASRNDYSYWLSTPALMNNAPITGR 81  
 DR 61 ALEPYISRCTVCEPGLATAIVHSQTFDIPPOPHGNLWKGFSFIMFTSAGSEGIGQALA 120  
 DR 82 ALEPYISRCTVCEPGLATAIVHSQTFDIPPOPHGNLWKGFSFIMFTSAGSEGIGQALA 141  
 DR 121 SPGSCLEEFRASPFLCHGRCTCNYSNSYSFWLASLNPERMFRKPIPSVYKAGELEKII 180  
 DR 142 SPGSCLEEFRASPFLCHGRCTCNYSNSYSFWLASLNPERMFRKPIPSVYKAGELEKII 201  
 QY 181 SRCQVCMKKRH 191  
 DB 202 SRCQVCMKKRH 212

DR EMBL: L47282; AAA91304..1; -.  
 DR InterPro; IPR00504; RRM.  
 DR InterPro; IPR001442; C4.  
 DR NCBI\_TAXID=9940;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-RENAL CORTEX;  
 RA Turner A.N., Ryan J.J., Derry C.J., Cashman S.J., Kattabma I.,  
 RA Mason P.J., Pusey C.D.;  
 RA Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: L47282; AAA91304..1; -.  
 DR InterPro; IPR00504; RRM.  
 DR InterPro; IPR001442; C4.







Qy 1 GLGKKGKDGSGPATWTRGFVFRHSSOTTAIPSCPEGVPLYSGFSFLFVQGNQRAHQD 60  
 |||.....|||.....|||.....|||.....|||.....|||.....|||.....|||.....|||.....|||.....  
 Db 1427 GLRKGRGDGSGPATWTRGFVFRHSSOTTAIPSCPEGVPLYSGFSFLFVQGNQRAHQD 1486

Qy 61 LGTLGSLCQRFITMPFLFCNVNDVCNFASRNDYSYWLSTPALMMAPIGRALEYVIS 120  
 |||.....|||.....|||.....|||.....|||.....|||.....|||.....|||.....|||.....|||.....  
 Db 1487 LGTLGSLCQRFITMPFLFCNVNDVCNFASRNDYSYWLSTPALMMAPIGRALEYVIS 1546

Qy 121 RCTVCEGPAIAI 132  
 |||.....|||.....|||.....|||.....|||.....|||.....|||.....|||.....|||.....|||.....  
 Db 1547 RCTVCEGPAIAI 1558

RESULT 2

B49736 collagen alpha 3(IV) chain, medium splice form - human (fragment)  
 Contains: collagen alpha 3(IV) chain, splice form GP-V

Species: Homo sapiens (man)  
 Date: 03-May-1994 #sequence\_revision 12-Nov-1999 #text\_change 17-Mar-2000

Accession: B49736; D49736; S60111

R. Feng, L.; Xia, Y.; Wilson, C.B.

J. Biol. Chem. 269: 2342-2348, 1994

A;Title: Alternative Splicing of the NCl domain of the human alpha3(IV) collagen gene.

A;Reference number: A49736; MUID:94124597

A;Status: nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 169-220 <FPN1>

A;Accession: D49736  
 A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: mRNA  
 A;Residues: 22-220 <FPN2>

A;Cross-references: GB:U02519; NID:9409106; PIDN:AAA18942.1; PID:9409107

A;Note: this is the conceptual translation of the nucleic acid submitted to GenBank

R. Pendas, J.R.; Bernal, D.; Revart, F.; Johansson, C.; Fresquet, V.J.; Cervera, J.; Wie

Eur. J. Biochem. 229: 754-760, 1995

A;Title: Characterization and expression of multiple alternatively spliced transcripts c

A;Antigen and one of its alternative forms.

A;Reference number: S69111; MUID:95278230

A;Accession: S69111

A;Molecule type: mRNA

A;Residues: 1-45,169-204,'L',206-220 <PEN>

C;Comment: For the complete sequence of the long splice form, see PIR:CGHU3B.

C;Genetics: GDB:COL4A3  
 A;Cross-references: GDB:128351; OMIM:120070

A;Gene: GDB:COL4A3  
 A;Position: 2q36-2q37

C;Peptide: collagen alpha 1(IV) chain  
 C;Keywords: alternative splicing; basement membrane; cell binding; coiled coil; extracellular

F;1-220;Product: collagen alpha 3(IV) chain, medium splice form (fragment) #status pred;  
 F;1-45,169-220;Product: collagen alpha 3(IV) chain, splice from GP-V (fragment) #status

F;22-220;Domain: carboxyl-terminal nonhelical, NCl <ncl>  
 F;34-134;Domain: collagen IV carboxyl-terminal repeat <ctn>

Query Match 98.9%; Score 713; DB 2; Length 220;  
 Best Local Similarity 99.2%; Pred. No. 1.4e-65; Mismatches 131; Conservative 0; Indels 0; Gaps 0;

Db 1 GLKKGKDGSGPATWTRGFVFRHSSOTTAIPSCPEGVPLYSGFSFLFVQGNQRAHQD 60

Qy 10 GLRKGRGDGSGPATWTRGFVFRHSSOTTAIPSCPEGVPLYSGFSFLFVQGNQRAHQD 69

Db 61 LGTLGSLCQRFITMPFLFCNVNDVCNFASRNDYSYWLSTPALMMAPIGRALEYVIS 120

Qy 70 LGTLGSLCQRFITMPFLFCNVNDVCNFASRNDYSYWLSTPALMMAPIGRALEYVIS 129

Qy 121 RCTVCEGPAIAI 132

Db 130 RCTVCEGPAIAI 141

Search completed: March 6, 2002, 06:54:49  
 Job time: 40 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:

March 6, 2002, 07:05:08 ; Search time 18.75 Seconds

(without alignments)  
258.120 Million cell updates/sec

Title: US-09-543-371-10\_COPY\_1\_132  
721 GLKQKRGDSGSPATWTRGF.....RALEPYVISRCTIVCEGPAIAI 132

Sequence: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 10059 seqs, 36664827 residues

Number of hits satisfying chosen parameters: 1

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 90%  
Maximum Match 100%

Listing first 1000 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query	Match Length	DB ID	Description
1	Q01955	721	CA34_HUMAN	homo sapien

#### ALIGNMENTS

RESULT	1	CA34_HUMAN	STANDARD:	PRT:	1670 AA.
T	001955;				
D	01-OCT-1996 (Rel. 34, Created)				
D	15-JUL-1999 (Rel. 38, Last sequence update)				
D	20-AUG-2001 (Rel. 40, Last annotation update)				
D	COLLAGEN ALPHA 3(IV) CHAIN PRECURSOR (GOODPASTURE ANTIGEN).				
GN	COL4A3.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
NCBI_TAXID=9606;					
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	Medline:9436894; PubMed=8003201;				
RX	Mariyama M., Leinonen A., Mochizuki T., Tryggvason K., Reenders S.T.;				
RT	Complete primary structure of the human alpha 3(IV) collagen chain.				
RT	Coxexpression of the alpha 3(IV) and alpha 4(IV) collagen chains in human tissues"; J. Biol. Chem. 269:23013-23017(1994).				
RT	[2]				
RP	REVISIONS.				
RA	J. Am. Soc. Nephrol. 12:97-105 (2001).				
RA	Leinonen A.;				
RL	Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.				

			CC
-	-	SUBUNIT: THERE ARE SIX TYPE IV COLLAGEN ISOFORMS, ALPHA 1(IV)-ALPHA 6(IV), EACH OF WHICH CAN FORM A TRIPLE HELIX STRUCTURE WITH 2 OTHER CHAINS TO GENERATE TYPE IV COLLAGEN NETWORK.	CC
CC	CC	SUBCELLULAR LOCATION: CELL SURFACE (POTENTIAL).	CC
CC	CC	- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS; 1 (SHOWN HERE), 2/4 AND 3/4/5, ARE PRODUCED BY ALTERNATIVE SPLICING. THEY DIFFER IN THEIR C-TERMINAL NCI DOMAINS.	CC
CC	CC	- TISSUE SPECIFICITY: ALPHA 3 AND ALPHA 4 TYPE IV COLLAGENS ARE COLOCALIZED AND PRESENT ONLY IN BASEMENT MEMBRANES OF KIDNEY, EYE, COCHLEA, LUNG AND BRAIN.	CC
CC	CC	- DOMAIN: ALPHA CHAINS OF TYPE IV COLLAGEN HAVE A NONCOLLAGENOUS DOMAIN (NCI) AT THEIR C-TERMINUS, FREQUENT INTERRUPTIONS OF THE G-X-Y REPEATS IN THE LONG CENTRAL TRIPLE HELICAL DOMAIN (WHICH MAY CAUSE FLEXIBILITY IN THE TRIPLE HELIX), AND A SHORT N-TERMINAL TRIPLE-HELICAL 7S DOMAIN.	CC
CC	CC	- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPETIDE REPEATING UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.	CC
CC	CC	- PTM: THE ALTERNATIVE SPliced FORM V CONTAINS AN ADDITIONAL N-LINKED GLYCOSYLATION SITE.	CC
CC	CC	- PTM: TYPE IV COLLAGENS CONTAIN NUMEROUS CYSTEINE RESIDUES WHICH ARE INVOLVED IN INTER- AND INTRAMOLECULAR DISULFIDE BONDING. 12 OF THESE, LOCATED IN THE NCI DOMAIN, ARE CONSERVED IN ALL KNOWN TYPE IV COLLAGENS.	CC
CC	CC	- DISEASE: ANTIBODIES AGAINST THE NCI DOMAIN OF ALPHA3(IV) MEDIATE THE HUMAN AUTOIMMUNE DISEASE, GOODPASTURE SYNDROME, WHICH IS CHARACTERIZED BY HEMATURIA AND PULMONARY HEMORRHAGE.	CC
CC	CC	- DISEASE: DEFECTS IN COL4A3 ARE ASSOCIATED WITH THE TYPE I AUTOSOMAL RECESSIVE FORM OF ALPORT SYNDROME, AN HEREDITARY GLOMERULONEPHROPATHY CHARACTERIZED BY PROGRESSIVE RENAL FAILURE, HEMATURIA AND DEAFNESS. THE RECESSIVE FORM OCCURS EQUALLY BETWEEN MALES AND FEMALES.	CC
CC	CC	-----	CC
CC	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation of the European Bioinformatics Institute. There are no restrictions on use by non profit institutions as long as its content is in no modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announcements.html">http://www.isb-sib.ch/announcements.html</a> or send an email to license@isb-sib.ch).	CC
CC	CC	-----	CC
DR	EMBL; X00031; CAA55335.1; -.		CC
DR	EMBL; M22993; AAA21610.1; -.		CC
DR	MIM; 203780; -.		CC
DR	EMBL; M55790; AAB19637.1; -.		CC
DR	EMBL; M81379; AAA1556.1; -.		CC
DR	EMBL; L08650; AAA52044.1; -.		CC
DR	EMBL; U02519; AAI18942.1; -.		CC
DR	EMBL; U02520; AAI18943.1; -.		CC
DR	EMBL; AB008495; BAM25064.1; -.		CC
DR	MIM; 120070; -.		CC
DR	EMBL; M22993; AAA21610.1; -.		CC
DR	InterPro; IPR01442; C4.		CC
DR	InterPro; IPR00087; collagen.		CC
DR	Pfam; PF01413; C4; 2.		CC
DR	Pfam; PF01391; collagen; 21.		CC
DR	ProDom; PD005923; C4; 2.		CC
DR	SMART; SM00111; C4; 2.		CC
KW	EXTRACELLULAR MATRIX; CONNECTIVE TISSUE; REPEAT; HYDROXYLATION; GLYCOPROTEIN; BASEMENT MEMBRANE; COLLAGEN; SIGNAL; CELL ADHESION; ALTERNATIVE SPLICING; POLYMORPHISM; PHOSPHORYLATION; DISEASE MUTATION; AIRPORT SYNDROME.		CC
KW	SIGNAL 1 28 POTENTIAL.		CC
FT	CHAIN 29 1670 COLLAGEN ALPHA 3(IV) CHAIN.		CC
FT	DOMAIN 29 42 7S DOMAIN.		CC
FT	DOMAIN 43 1438 TRIPLE-HELICAL REGION.		CC
FT	DOMAIN 1439 1670 NONHELCICAL REGION (NCI). (GOODPASTURE ANTIGEN) (BY SIMILARITY). EPITOPE (RECOGNIZED BY GOODPASTURE ANTIBODIES).		CC
FT	DOMAIN 1427 1444 CLEAVAGE (BY COLLAGENASE) (BY SIMILARITY).		CC
FT	SITE 1426 1427 CELL ATTACHMENT SITE (POTENTIAL).		CC
FT	SITE 791 793		CC

FT	SITE	996	998	CELL ATTACHMENT SITE (POTENTIAL).
FT	SITE	1154	1156	CELL ATTACHMENT SITE (POTENTIAL).
FT	SITE	1308	1345	CELL ATTACHMENT SITE (POTENTIAL).
FT	SITE	1345	1347	CELL ATTACHMENT SITE (POTENTIAL).
FT	SITE	1434	1432	CELL ATTACHMENT SITE (POTENTIAL).
FT	CARBONYD	253	253	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	MOD-RES	1435	1435	PHOSPHORYLATION (BY SIMILARITY).
FT	MOD-RES	1437	1437	PHOSPHORYLATION (BY SIMILARITY).
FT	DISUFTID	1460	1551	OR 1548 (BY SIMILARITY).
FT	DISUFTID	1493	1548	OR 1551 (BY SIMILARITY).
FT	DISUFTID	1505	1511	BY SIMILARITY.
FT	DISUFTID	1570	1665	OR 1662 (BY SIMILARITY).
FT	DISUFTID	1604	1662	OR 1665 (BY SIMILARITY).
FT	DISUFTID	1616	1622	BY SIMILARITY.
FT	VARSPLIC	1586	1670	FTSAGSEGCGOALASPGSLEEFRAFPFLFCHRGTCNYNSYSFWLASLNSPERMFRRKIPSPVKAEGELEKJITSRQCVCKMKNRKH -> KAVSYINCEWSGIRKKNNLSGHWEEHKTLKUKTAELVFTENKVKYTMVHAVI (IN ISOFORM 2).
FT	VARSPLIC	1488	1670	GTLGSCQLQRWTTMPFLCFNWDVNCFAFRNDISYMLSPALMPMMNAPITGRALEPYBVISRCTVPEGATAVASQTDIPPCHPMHSIANGSEFMTGSGSAGSSEGCGOALASPGSLEEFRAFPFLFCHRGTCNYNSYSFWLASLNSPERMFRRKIPSPVKAEGELEKJITSRQCVCKMKNRKH -> DALFVYKVLRSP (IN ISOFORM 3).
FT	VARIANT	43	43	G -> R.
FT	VARIANT	162	162	/FTId=VAR_011202.
FT	VARIANT	297	297	G -> E.
FT	VARIANT			/FTId=VAR_011203.
FT	VARIANT			G -> E (IN AS).
Query	Match	100 %;	Score	721;
Best	Local Similarity	100 %;	DB	1;
Matches	132;	Conservative	No.	7_9e-68;
Oy		0;	Mismatches	0;
Db		0;	Indels	0;
Oy		0;	Gaps	0;
Db		0;		0;
Query	61	LGTLSCLQRTTMFLCWNDVCNEASRHYWLSUPRALPMMARITGAEPVIS	120	
Db	1487	LGTLSCLQRTTMFLFCNVNDVCNEASRNDSYWLSTPALPMNMAPITGRALEPVIS	1486	
Oy	121	RCTVCEGPATAI	132	
Db	1547	RCTVCEGPATAI	1558	

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: March 6, 2002, 07:04:44 ; Search time 50.56 Seconds  
(without alignments)

US-09-543-371-10\_COPY\_1\_132  
381.882 Million cell updates/sec

Title: US-09-543-371-10\_COPY\_1\_132  
perfect score: 721  
Sequence: 1 GLKGKRGDGSQSPATWTRGF.....RALEPYISRCVCEGPAIAI 132  
Scoring table: BLOSUM62  
Gapov 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

To number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 90%  
Maximum Match 100%  
Listing first 1000 summaries

Database : SPREMBL\_17:\*

- 1: sp\_archea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mic:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result	Query	Score	Match	Length	DB	ID	Description
1	721	100.0	245	4	09NYC4		09nyc4 homo sapien
2	721	100.0	1670	4	09BQT2		09bqt2 homo sapien

## ALIGNMENTS

RESULT	1
09NYC4	PRELIMINARY; PRT; 245 AA.
AC	09NYC4;
DT	01-OCT-2000 (TREMBLrel. 15, Created) 01-OCT-2000 (TREMBLrel. 15, Last sequence update) 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DR	01-JUN-2001 (TREMBLrel. 17, Last annotation update) ALPHA3 TYPE IV COLLAGEN. COLA3. Homo sapiens (Human). Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. NCBI_TaxID=9606;
RN	[1] SEQUENCE FROM N.A. RP MEDLINE=21064695; PubMed=11134255;
RA	Heidet L., Arrendel C., Forestier L., Cohen-Solal L., Mollet G., Gutierrez B., Savrou C., Gubler M.C., Antignac C., "Structure of the human type IV collagen gene COLA3 and mutations in autosomal Alport syndrome," J. Am. Soc. Nephrol. 12:97-106 (2001).
RT	J. Am. Soc. Nephrol. 12:97-106 (2001).
RL	EMBL; AJ288487; CAC36101.1; JOINED.
DR	EMBL; AJ288488; CAC36101.1; JOINED. EMBL; AJ288489; CAC36101.1; JOINED. EMBL; AJ288490; CAC36101.1; JOINED. EMBL; AJ288491; CAC36101.1; JOINED. EMBL; AJ288492; CAC36101.1; JOINED. EMBL; AJ288493; CAC36101.1; JOINED. EMBL; AJ288494; CAC36101.1; JOINED. EMBL; AJ288495; CAC36101.1; JOINED. EMBL; AJ288496; CAC36101.1; JOINED. EMBL; AJ288497; CAC36101.1; JOINED. EMBL; AJ288498; CAC36101.1; JOINED. EMBL; AJ288499; CAC36101.1; JOINED. EMBL; AJ288501; CAC36101.1; JOINED. EMBL; AJ288502; CAC36101.1; JOINED. EMBL; AJ288503; CAC36101.1; JOINED.

DR EMBL; AJ288504; CAC36101.1; JOINED.  
 DR EMBL; AJ288505; CAC36101.1; JOINED.  
 DR EMBL; AJ288506; CAC36101.1; JOINED.  
 DR EMBL; AJ288507; CAC36101.1; JOINED.  
 DR EMBL; AJ288508; CAC36101.1; JOINED.  
 DR EMBL; AJ288509; CAC36101.1; JOINED.  
 DR EMBL; AJ288510; CAC36101.1; JOINED.  
 DR EMBL; AJ288511; CAC36101.1; JOINED.  
 DR EMBL; AJ288512; CAC36101.1; JOINED.  
 DR EMBL; AJ288513; CAC36101.1; JOINED.  
 DR EMBL; AJ288514; CAC36101.1; JOINED.  
 DR EMBL; AJ288515; CAC36101.1; JOINED.  
 DR EMBL; AJ288516; CAC36101.1; JOINED.  
 DR EMBL; AJ288517; CAC36101.1; JOINED.  
 DR EMBL; AJ288518; CAC36101.1; JOINED.  
 DR EMBL; AJ288519; CAC36101.1; JOINED.  
 DR EMBL; AJ288520; CAC36101.1; JOINED.  
 DR EMBL; AJ288521; CAC36101.1; JOINED.  
 DR EMBL; AJ288522; CAC36101.1; JOINED.  
 DR EMBL; AJ288523; CAC36101.1; JOINED.  
 DR EMBL; AJ288524; CAC36101.1; JOINED.  
 DR EMBL; AJ288525; CAC36101.1; JOINED.  
 DR EMBL; AJ288526; CAC36101.1; JOINED.  
 DR EMBL; AJ288527; CAC36101.1; JOINED.  
 DR EMBL; AJ288528; CAC36101.1; JOINED.  
 DR EMBL; AJ288529; CAC36101.1; JOINED.  
 DR EMBL; AJ288530; CAC36101.1; JOINED.  
 DR EMBL; AJ288531; CAC36101.1; JOINED.  
 DR EMBL; AJ288532; CAC36101.1; JOINED.  
 DR EMBL; AJ288533; CAC36101.1; JOINED.  
 DR EMBL; AJ288534; CAC36101.1; JOINED.  
 DR EMBL; AJ288535; CAC36101.1; JOINED.  
 DR EMBL; AJ288536; CAC36101.1; JOINED.  
 DR EMBL; AJ288537; CAC36101.1; JOINED.  
 DR EMBL; AJ288538; CAC36101.1; JOINED.

SEQUENCE 1670 AA; 161899 MW; FA7BB4914CA0A6FF CRC64;

Query Match 100.0%; Score 721; DB 4; Length 1670;  
 Best Local Similarity 100.0%; Pred. No. 1 9e-71;  
 Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GIGKKGKRGDSGSPATWTRGFVTRHSQTTAIPSCPEGTYPLYSGFSEFLFGQNRRAHGQD 60
Db	1427	GLKGKRGDSGSPATWTRGFVTRHSQTTAIPSCPEGTYPLYSGFSEFLFGQNRRAHGQD 1486
	61	LGTIGLGSCLQRFITMPFELCNVNNDVCNFASRADSYWILSTPALPMNMAPITGRALEPYIS 120
Db	1487	LGTLGSLCQRFITMPFELCNVNNDVCNFASRDYSYMSTPALPMNMAPITGRALEPYIS 1546
QY	121	RCTVCEGPAlI 132
Db	1547	RCTVCEGPAlI 1558



DT 03-APR-2000 (first entry)  
 XX DE Human type IV collagen alpha 3 chain protein sequence SEQ ID NO:10.  
 XX KW Human; type IV collagen; anti-angiogenesis; angiogenesis; cancer;  
 KW ocular angiogenesis; diabetic retinopathy; psoriasis;  
 KW myocardial angiogenesis; plaque neovascularisation; angiofibroma;  
 KW atherosclerosis; scleroderma; hypertrophic scar; cat scratch disease;  
 KW contraception; obesity.  
 OS Homo sapiens.  
 XX PN WO9965940-A1.  
 XX PD 23-DEC-1999.  
 XX PT 17-JUN-1999; 99WO-US13737.  
 XX 17-JUN-1998; 98US-0089689.  
 XX 25-MAR-1999; 99US-0126175.  
 XX PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.  
 XX PT Kalluri, R.;  
 XX DR N-PSDB; AA257158.  
 XX PT Anti-angiogenic Proteins comprising the NC1 domain of the alpha 1, 2 or  
 3 chain of Type IV collagen used in, e.g. treatment of benign tumors  
 PT and rheumatoid arthritis -  
 XX Claim 32; Fig 16B; 117pp; English.  
 XX  
 CC The present sequence represents the human type IV collagen alpha 3 chain.  
 CC The present invention describes an isolated protein chosen from the NC1  
 CC domain of the alpha 1, alpha 2 or alpha 3 chains of type IV collagen or  
 CC a fragment, analogue, derivative or mutant, which has anti-angiogenic  
 properties. The anti-angiogenic proteins, multimers and chimeras are  
 CC especially useful for inhibiting angiogenic activity in mammalian tissue,  
 CC cancers, benign tumours, rheumatoid arthritis, diabetic retinopathy,  
 CC psoriasis, ocular angiogenesis, Osler-Weber Syndrome,  
 CC myocardial angiogenesis, plaque neovascularisation, telangiectasia,  
 CC haemophilic joints, angiofibroma, wound granulation, intestinal  
 CC adhesions, atherosclerosis, scleroderma, hypertrophic scars, cat scratch  
 CC disease, Helicobacter pylori ulcers, dialysis graft vascular access  
 CC stenosis, contraction and obesity. The compositions can be used to  
 CC inhibit a disease characterised by angiogenic activity, in conjunction  
 CC with radiation therapy, chemotherapy or immunotherapy.  
 XX Sequence 245 AA;  
 XX  
 Query Match 100.0%; Score 721; DB 21; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-72;  
 Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 GLKCKRGKGSGSPARWTTTRGFVFTFRHSOTTAIPSCPEGTYPLYSGFSFLFGQGNQRAGQD 60  
 Db 2 gIgkrgkgsgspatwttrgfvtfrhsottaipscpegtyplysgfsflfgqnrabqgd 61.  
 OY 61 LGTIGSGCQRFATMPFLCNVNDFASRNDDSYWLSTPALMMAPNMAPITGEALEPYIS 120  
 Db 62 lgltgsciqrfatmpfifcnvnvdnfasrndsywlstpalmmapnmapitalepyis 121  
 OY 121 RCTVCEGAIAT 132  
 Db 122 rctvcegpaia 133  
 CC RESULT 3

AAY31993 ID AAY31993 standard; Protein: 268 AA.  
 XX AC AAY31993;  
 XX DE DT 05-JAN-2000 (first entry)  
 XX KW Type IV collagen NC1 domain alpha-3 monomer.  
 KW Type IV collagen; NC1 domain; non-collagenous domain; human;  
 KW angiogenesis; tumour; metastasis; therapy; diabetic retinopathy;  
 KW rheumatoid arthritis; retinal neovascularization; macular degeneration;  
 KW choroidal neovascularization; retinopathy of prematurity;  
 KW corneal graft rejection; neovascular glaucoma;  
 KW retrobulbar fibroplasia; epidemic keratoconjunctivitis;  
 KW vitamin A deficiency; contact lens overwear; atopic keratitis;  
 KW superior limbic keratitis; pterygium keratitis sicca; sogrens;  
 KW acne rosacea; phylectenullosis; syphillis; Mycobacteria infection;  
 KW lipid degeneration; chemical burn; ulcer; herpes simplex infection;  
 KW Herpes zoster infection; protozoan infection; Kaposi's sarcoma;  
 KW Mooren ulcer; Terrien's marginal degeneration;  
 KW marginal keratolysis; trauma; systemic lupus; polyarteritis;  
 KW Wegener's sarcoidosis; scleritis; Steven's Johnson disease;  
 KW radial keratopathy; sickle cell anemia; sarcoid;  
 KW pseudoxanthoma elasticum; Paget's disease; vein occlusion;  
 KW artery occlusion; carotid obstructive disease; chronic uveitis;  
 KW myopia; optic pit; Starckert's disease; Eales disease; Bechets disease;  
 KW chronic retinal detachment; fibrovascular tissue proliferation;  
 KW haemangioma; osler-weber-rendu; ARDS; ocular neovascular disease;  
 KW osteoarthritis; chronic inflammation; Crohn's disease;  
 KW ulcerative colitis; psoriasis; atherosclerosis; pemphigoid.  
 XX OS Homo sapiens.  
 OS Synthetic.  
 XX FH Key Location/Qualifiers  
 FT Peptide 1..17  
 FT Protein /note= "BM40 signal peptide"  
 FT Peptide 18..268  
 FT Protein /note= "mature protein"  
 FT Protein 18..25  
 FT Protein /note= "affinity tag"  
 FT Protein 25..268  
 FT Protein /note= "NC1 alpha-3 monomer"  
 XX PN WO9949885-A2.  
 XX PD 07-OCT-1999.  
 XX PT 26-MAR-1999; 99WO-US06445.  
 XX PR 27-MAR-1998; 98US-0079783.  
 XX PR 29-OCT-1998; 98US-0106170.  
 XX PA (UNIV ) UNIV KANSAS MEDICAL CENT.  
 XX PI Hudson BG, Sarras MP;  
 XX DR WPI; 1999-601297/51.  
 XX N-PSDB; AA220091.  
 XX PT Inhibition of angiogenesis with non-collagenous alpha chain monomer  
 PT useful for treating e.g. tumor growth or metastasis,  
 PT neovascularisation, etc.  
 XX Disclosure; Fig 17c; 56pp; English.  
 XX This sequence represents a recombinant type IV collagen  
 CC non-collagenous (NC1) domain alpha-3 polypeptide composed of a  
 BM40 signal sequence (which is cleaved from the mature protein) to

CC facilitate protein secretion, and a mature protein comprising an affinity tag (facilitates purification and identification of the material) and the alpha-1 chain monomer. The invention provides methods and kits for inhibiting angiogenesis, tumour growth and metastasis, and endothelial cell interaction with the extracellular matrix, each method comprising contacting the tumour or animal tissue with 1 or more isolated type IV collagen NC1 alpha chain monomer(s) selected from the group consisting of alpha-1, alpha-2, alpha-3 and alpha-6 NC1 chain monomers (see AAY31991-96). The monomers can be produced via recombinant protein expression. The polynucleotides and polypeptides are used to treat an angiogenesis-mediated disorder or condition, especially selected from solid and blood-borne tumours, diabetic retinopathy, rheumatoid arthritis, retinal neovascularization, choroidal neovascularization, macular degeneration, corneal graft rejection, neovascular glaucoma, retrobulbar fibroplasia, epidemic keratoconjunctivitis, vitamin A deficiency, contact lens overwear, atopic keratitis, superior limbic keratitis, pterygium keratitis sicca, sogns, acne rosacea, phylogenesis, chemical burns, bacterial ulcers, fungal ulcers, herpes simplex infections, herpes zoster infections, protozoan infections, Kaposi's sarcoma, Mooren ulcer, Terrien's marginal degeneration, marginal keratolysis, trauma, systemic lupus, polyarteritis, Wegener's sarcoidosis, scleritis, Steven's Johnson disease, radial keratotomy, sickle cell anaemia, sarcoid, pseudoxanthoma elasticum, Paget's disease, vein occlusion, artery occlusion, carotid obstructive disease, chronic uveitis, chronic vitritis, Lyme's disease, Eales disease, Bechets disease, myopia, optic pits, Stargarts disease, pars planitis, chronic retinal detachment, hydropsis, toxicoplasmosis, post-laser complications, abnormal proliferation of fibrovascular tissue, haemangiomas, Osler-Weber-Rendu, AIDS, ocular neovascular disease, osteoarthritis, chronic inflammation, Crohn's disease, ulcerative colitis, psoriasis, atherosclerosis, and pemphigoid (all claimed).

XX Sequence 268 AA;

Query Match 97.1%; Score 700; DB 20; Length 268;  
Best Local Similarity 100.0%; Pred. No. 4.8e-70; Mismatches 0; Indels 0; Gaps 0;

Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC Qy 5 KRGDSSGSPATWTRGVFTRHSQTTAIPSCPEGTPLVLYSGFSFLFVQGNQRAGHDGTL 64  
Db 29 krgdsgspatwtrgvftrhsqttalpscpegtplvlysgfsflfvqgnqrhggdgtl 88

Qy 65 GSCLQRTTMLFCNVNDVNFASRNDYSWLSTPALMPNMAPITGRALEPIYSRCTV 124  
D 89 gsclqrftttmpfifcnvdvfasrndyswvlstpalmppnmapitgralepyisrcrv 148

Qy 125 CEGPAIAI 132  
Db 149 cegpaiai 156

RESULT 4  
ID AAY97555 standard; Protein: 268 AA.  
XX AC AAY97555;  
AC AAY97555;  
DE Human alpha3(IV)NC1 protein sequence.

XX KW chronic inflammation; psoriasis; therapy; alpha3(IV)NC1.  
KW XX OS Homo sapiens.  
XX PN WO20059532-A1.  
XX PD 12-OCT-2000.  
XX PF 31-MAR-2000; 2000WO-US08678.  
XX PR 01-APR-1999; 99US-0127391.  
XX PA (BIOS-) BIOSTRUCTURE INC.  
XX PI Brooks P, Hudson B;  
XX DR WPI; 2000-664962/64.  
XX N-PSDB; AAY90993.

SQ Use of antagonists of specific integrin receptors for inhibiting angiogenesis, tumour growth or metastases, or endothelial cell interactions with the extracellular matrix Disclosure; Fig 17c; 78pp; English.

CC This sequence is a human type IV collagen alpha chain monomer, designated alpha3(IV)NC1. The invention relates to a method for inhibiting angiogenesis, tumour growth or metastases, or endothelial cell interactions with the extracellular matrix, comprising contacting the cells or tissue with a polypeptide composition containing antagonists of specific integrin receptors. The methods and the antagonists are useful for inhibiting angiogenesis, tumour growth or metastases, or endothelial cell interaction with the extracellular matrix. The antagonists are also useful for treating diseases and conditions with accompanying undesired angiogenesis, e.g. solid and blood-borne tumours (e.g. melanomas, carcinomas, sarcomas, rhabdomyosarcoma, retinoblastoma, Ewing sarcoma, neuroblastoma, osteosarcoma or leukemia). These are also applicable to treating non-tumorigenic diseases and conditions with accompanying undesired angiogenesis, e.g. diabetic retinopathy, rheumatoid arthritis, retinal neovascularisation, choroidal neovascularisation, muscular degeneration, corneal graft rejection, vitamin A deficiency, atopic keratitis, Mycobacteria infections, chemical burns, Kaposi's sarcoma, sickle cell anaemia, sarcoid, carotid obstructive disease, post-laser complications, chronic inflammation or psoriasis.

XX Sequence 268 AA;

Query Match 97.1%; Score 700; DB 21; Length 268;  
Best Local Similarity 100.0%; Pred. No. 4.8e-70; Mismatches 0; Indels 0; Gaps 0;

Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC Qy 5 KRGDSSGSPATWTRGVFTRHSQTTAIPSCPEGTPLVLYSGFSFLFVQGNQRAGHDGTL 64  
Db 29 krgdsgspatwtrgvftrhsqttalpscpegtplvlysgfsflfvqgnqrhggdgtl 88

Qy 65 GSCLQRTTMLFCNVNDVNFASRNDYSWLSTPALMPNMAPITGRALEPIYSRCTV 124  
D 89 gsclqrftttmpfifcnvdvfasrndyswvlstpalmppnmapitgralepyisrcrv 148

Qy 125 CEGPAIAI 132  
Db 149 cegpaiai 156

Search completed: March 6, 2002, 06:56:24  
Job time: 135 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on:

March 6, 2002, 06:55:22 ; Search time 26.9 Seconds

(without alignments)  
110.425 Million cell updates/sec

Title: US-09-543-371-10\_COPY\_1\_132

Perfect score: 721

Sequence: 1 GLKGKRGDGGSPATWTTGCF.....RALEPYIISRCTVCEGPAIAI 132

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 90%

Maximum Match 100%

Listing first 1000 summaries

Database :

Issued\_Patents\_AA:\*

```

1: /cgn2_5/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_5/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_5/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_5/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_5/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_5/ptodata/2/iaa/backfilesl.pep:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query Score	Match Length	DB ID	Description

No matches found

Search completed: March 6, 2002, 06:55:22  
Job time: 73 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: March 6, 2002, 06:54:49 ; Search time 31.16 Seconds

(without alignments)  
156,456 Million cell updates/sec

Title: US-09-543-371-10\_COPY\_181\_244

Perfect score: 353

Sequence: 1 EFRASPFLECHGRGTCNVYS..... KAGELEKLISRQVCMKKRRH 64

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 3

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 90%  
Maximum Match 100%

Listing first 1000 summaries

Database : PIR:68;\*

1: pir1;\*  
2: pir2;\*  
3: pir3;\*  
4: pir4;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	353	100.0	1670	1 CGHU3B	collagen alpha 3(I)
2	345	97.7	246	2 I48302	collagen alpha 3(I)
3	323	91.5	471	2 A39024	collagen alpha 3(I)

## ALIGNMENTS

RESULT 1  
CGHU3B  
collagen alpha 3(IV) chain precursor, long splice form - human  
C:Alternate names: Goodpasture antigen; procollagen alpha 3(IV) chain long splice form  
C:Species: Homo sapiens (man)  
C:Date: 28-Oct-1994 #sequence\_revision 03-Oct-1995 #text\_change 22-Jun-1999  
C:Accession: A54763; A43928; A40434; A45971; A39786  
R:Mariyama, M.; Leinonen, A.; Mochizuki, T.; Tryggrason, K.; Reeder, S.T.  
J. Biol. Chem. 269, 23013-23017, 1994  
A:Title: Primary structure of the human alpha3(IV) collagen chain. Coexpression  
A:Reference number: A54763; MUID:94364994  
A:Accession: A54763  
A:Molecule type: mRNA  
A:Residues: 1331-1524 'I', 1526-1670 <TUR>  
A:Cross-references: GR: M81379  
A:Experimental source: kidney  
R:Quinones, S.; Bernal, D.; Garcia-Sogo, M.; Elena, S.F.; Saus, J.  
J. Biol. Chem. 267, 19780-19784, 1992  
A:Title: Exon/intron structure of the human alpha 3(IV) gene encompassing the Goodpasture antigen.  
A:Reference number: A44043; MUID:93015826  
A:Cross-references: GR: M92993; NID:9177835; PIDN:AA21610.1; PID:9177896  
R:Quinones, S.; Bernal, D.; Garcia-Sogo, M.; Elena, S.F.; Saus, J.  
J. Biol. Chem. 269, 17358, 1994  
A:Reference number: A44738; MUID:9427434  
A:Content: annotation; erratum; correction to intronic sequence in A44043  
R:Bernal, D.; Quinones, S.; Saus, J.  
J. Biol. Chem. 268, 12090-12094, 1993  
A:Title: The human mRNA encoding the Goodpasture antigen is alternatively spliced.  
A:Reference number: A45971; MUID:93280184  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1427-1444 <BER>  
A:Note: sequence extracted from NCBI backbone (NCBIP:133363); sequence incorrectly id  
R:Morrison, K.E.; Mariyama, M.; Yang-Feng, T.L.; Reeders, S.T.  
Am. J. Hum. Genet. 49, 545-554, 1991  
A:Title: Sequence and localization of a partial cDNA encoding the human alpha3 chain  
A:Reference number: A39786; MUID:91353570  
A:Accession: A39786  
A:Molecule type: mRNA  
A:Residues: 1453-1593 'A', 1595-1670 <MR>  
A:Cross-references: GB: M53790; NID:9234418; PIDN:AB19637.1; PID:9234419  
C:Comment: In Goodpasture's syndrome, an autoimmune response develops against an epitope and subsequently O-glycosylated.  
C:Genetics:

A:Gene: GDB:COL4A3  
A:Cross-references: GDB:128351; OMIM:120070  
A:Map position: 1q36-q37  
A:Interrns: 1385Y1; 1418Y1; 1488Y1; 1547Y2; 1585Y3; 1643Y2 #status incomplete  
A:Note: the alpha 3(IV) and alpha 4(IV) chain genes are encoded on opposite strands w  
C:Complex: This minor type IV collagen is thought to form a heterotrimer of two alpha  
meng trimer amino-terminal domains (with disulfide and desmosine cross-links), dimeri  
er associations in the interrupted helical domain (with disulfide and desmosine cross  
C:Function:

A:Description: minor structural component of extracellular basement membrane in kidney  
C:Superfamily: collagen alpha 1(IV) chain  
C:Keywords: alternative splicing; basement membrane; cell binding; coiled coil; extra  
F:1-28/Domain: signal sequence  
F:1-28/Product: collagen alpha 3(IV) chain long splice form #status predicted <MA  
F:29-1670/Domain: amino-terminal nonhelical, NHI <NHI>  
F:29-42/Domain: amino-terminal interrupted helical  
F:43-1434/Region: cell attachment (R-G-D) motif  
F:996-998/Region: cell attachment (R-G-D) motif  
F:1154-1156/Region: cell attachment (R-G-D) motif  
F:1306-1308/Region: cell attachment (R-G-D) motif  
F:1345-1347/Region: cell attachment (R-G-D) motif  
F:1432-1434/Region: cell attachment (R-G-D) motif  
F:991-1000/Region: cell attachment (R-G-D) motif  
F:1451-1551/Domain: collagen IV carboxyl-terminal repeat <CT1>  
F:1561-1665/Domain: collagen IV carboxyl-terminal repeat <CT2>  
F:31-33, 39-41, 125-422, 476-479, 602, 722, 809-1387/Disulfide bonds: interchain #status predicted  
F:253/Binding site: carbohydrate-disulfide bonds: (or 1460-1551, 1493-1548) #status predicted  
F:1460-1548, 1493-1551/Disulfide bonds: (or 1460-1551, 1493-1548) #status predicted  
F:1570-1662, 1604-1665/Disulfide bonds: (or 1570-1662, 1604-1665) #status predicted  
A:Reference number: A43928; MUID:92147878  
A:Accession: A43928

Query Match 100.0%; Score 353; DB 1; Length 1670;  
Best Local Similarity 100.0%; Pred. No. 5.5e-35;

**RESULT 2**

QY 1 EFRASPFLCHGRTCNVYNSNSFWLASLNPERMFRKPIPSVKAGELEKITSRCQVM 60  
 C;Species: Mus musculus (house mouse)  
 C;Date: 15-Mar-1996 #sequence\_revision 15-Mar-1996 #text\_change 16-Feb-1997  
 C;Accession: I48302; S4728

R;Miner, J.H.; Saus, J.R.  
 J; Cell Biol. 127, 879-891, 1994  
 A;Title: Collagen IV alpha<sub>3</sub>, alpha<sub>4</sub>, and alpha<sub>5</sub> chains in rodent basal laminae: sequence  
 A;Accession: I48302  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-246 <RES>  
 A;Cross-references: EMBL:Z35166; NID:9535197; PID:9535198  
 C;Superfamily: collagen alpha 1(IV) chain

Query Match 97.7%; Score 345; DB 2; Length 246;  
 Best Local Similarity 95.3%; Pred. No. 7.8e-35; Indels 0; Gaps 0;  
 Matches 61; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EFRASPFLCHGRTCNVYNSNSFWLASLNPERMFRKPIPSVKAGELEKITSRCQVM 60  
 C;Accession: A39024; S20672; S17802; A35167; C39919; S1374; S20815  
 C;Author: Morrison, K.E.; Germino, G.G.; Reenders, S.T.  
 J; Biol. Chem. 266, 34-39, 1991  
 A;Title: Use of the polymerase chain reaction to clone and sequence a cDNA encoding the  
 A;Reference number: A39024; MUID:91093146  
 A;Accession: A39024  
 A;Molecule type: mRNA  
 A;Residues: 1-471 <NOR>  
 A;Cross-references: EMBL:M63139; NID:9162886; PID:AA62708.1; PID:9162887  
 R;Burkowski, R.J.; Langeweld, J.P.M.; Wieslander, J.; Hamilton, J.; Hudson, B.G.  
 J; Biol. Chem. 262, 7874-7877, 1987  
 A;Title: Localization of the Goodpasture epitope to a novel chain of basement membrane  
 A;Reference number: S18412; MUID:8722219  
 A;Accession: S20572  
 A;Molecule type: protein  
 A;Residues: 227-228, 'X', 230-244 <BUT>  
 R;Saus, J.; Wieslander, J.; Langeweld, J.P.M.; Quinones, S.; Hudson, B.G.  
 J; Biol. Chem. 263, 13380, 1988  
 A;Title: Identification of the Goodpasture antigen as the alpha<sub>3</sub>(IV) chain of collagen  
 A;Reference number: S17802; MUID:88330844  
 A;Accession: S17802  
 A;Molecule type: protein  
 A;Residues: 227-228, 'X', 230-252, 'Y', 254 <SAU>  
 R;Gunwar, S.; Saus, J.; Noelker, M.E.; Hudson, B.G.  
 J; Biol. Chem. 265, 5466-5469, 1990  
 A;Title: Glomerular basement membrane. Identification of a fourth chain, alpha<sub>4</sub>, of type

**RESULT 3**

A39024  
 collagen alpha 3(IV) chain - bovine (fragment)  
 C;Species: Bos primigenius taurus (cattle)  
 C;Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 13-Aug-1999  
 C;Accession: A39024; S20672; S17802; A35167; C39919; S1374; S20815  
 C;Author: Morrison, K.E.; Germino, G.G.; Reenders, S.T.  
 J; Biol. Chem. 266, 34-39, 1991  
 A;Title: Use of the polymerase chain reaction to clone and sequence a cDNA encoding the  
 A;Reference number: A39024; MUID:91093146  
 A;Accession: A39024  
 A;Molecule type: mRNA  
 A;Residues: 1-471 <NOR>  
 A;Cross-references: EMBL:M63139; NID:9162886; PID:AA62708.1; PID:9162887  
 R;Burkowski, R.J.; Langeweld, J.P.M.; Wieslander, J.; Hamilton, J.; Hudson, B.G.  
 J; Biol. Chem. 262, 7874-7877, 1987  
 A;Title: Localization of the Goodpasture epitope to a novel chain of basement membrane  
 A;Reference number: S18412; MUID:8722219  
 A;Accession: S20572  
 A;Molecule type: protein  
 A;Residues: 227-228, 'X', 230-244 <BUT>  
 R;Saus, J.; Wieslander, J.; Langeweld, J.P.M.; Quinones, S.; Hudson, B.G.  
 J; Biol. Chem. 263, 13380, 1988  
 A;Title: Identification of the Goodpasture antigen as the alpha<sub>3</sub>(IV) chain of collagen  
 A;Reference number: S17802; MUID:88330844  
 A;Accession: S17802  
 A;Molecule type: protein  
 A;Residues: 227-228, 'X', 230-252, 'Y', 254 <SAU>  
 R;Gunwar, S.; Saus, J.; Noelker, M.E.; Hudson, B.G.  
 J; Biol. Chem. 265, 5466-5469, 1990  
 A;Title: Glomerular basement membrane. Identification of a fourth chain, alpha<sub>4</sub>, of type

A;Reference number: A35167; MUID:90202779  
 A;Accession: A35167  
 A;Molecule type: protein  
 A;Residues: 236-258 <UNP>  
 R;Gunwar, S.; Ballester, F.; Kalluri, R.; Timoneda, J.; Chonko, A.M.; Edwards, S.J.;  
 J; Biol. Chem. 266, 15318-15324, 1991  
 A;Title: Glomerular basement membrane. Identification of dimeric subunits of the nonc  
 A;Reference number: A39419; MUID:91332055  
 A;Accession: C39419  
 A;Molecule type: protein  
 A;Residues: 236-255 <GL2>  
 C;Superfamily: collagen alpha 1(IV) chain  
 C;Keywords: basement membrane; cell binding; coiled coil; disulfide bond; duplication  
 F; 239-421/Domain: collagenous (fragment); #status predicted <COL>  
 F; 239-353/Domain: repeat NCL #status predicted <NC1>  
 F; 354-471/Domain: repeat NCL #status predicted <NC12>  
 F; 232,238/Modified site: hydroxyproline (Pro) #status experimental  
 F; 306-312,417-423/Disulfide bonds: #status predicted

Query Match 91.5%; Score 323; DB 2; Length 471;  
 Best Local Similarity 92.1%; Pred. No. 7.5e-32; Indels 0; Gaps 0;  
 Matches 58; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EFRASPFLCHGRTCNVYNSNSFWLASLNPERMFRKPIPSVKAGELEKITSRCQVM 60  
 C;Accession: A39024; S20672; S17802; A35167; C39919; S1374; S20815  
 C;Author: Morrison, K.E.; Germino, G.G.; Reenders, S.T.  
 J; Biol. Chem. 266, 34-39, 1991  
 A;Title: Use of the polymerase chain reaction to clone and sequence a cDNA encoding the  
 A;Reference number: A39024; MUID:91093146  
 A;Accession: A39024  
 A;Molecule type: mRNA  
 A;Residues: 1-471 <NOR>  
 A;Cross-references: EMBL:M63139; NID:9162886; PID:AA62708.1; PID:9162887  
 R;Burkowski, R.J.; Langeweld, J.P.M.; Wieslander, J.; Hamilton, J.; Hudson, B.G.  
 J; Biol. Chem. 262, 7874-7877, 1987  
 A;Title: Localization of the Goodpasture epitope to a novel chain of basement membrane  
 A;Reference number: S18412; MUID:8722219  
 A;Accession: S20572  
 A;Molecule type: protein  
 A;Residues: 227-228, 'X', 230-244 <BUT>  
 R;Saus, J.; Wieslander, J.; Langeweld, J.P.M.; Quinones, S.; Hudson, B.G.  
 J; Biol. Chem. 263, 13380, 1988  
 A;Title: Identification of the Goodpasture antigen as the alpha<sub>3</sub>(IV) chain of collagen  
 A;Reference number: S17802; MUID:88330844  
 A;Accession: S17802  
 A;Molecule type: protein  
 A;Residues: 227-228, 'X', 230-252, 'Y', 254 <SAU>  
 R;Gunwar, S.; Saus, J.; Noelker, M.E.; Hudson, B.G.  
 J; Biol. Chem. 265, 5466-5469, 1990  
 A;Title: Glomerular basement membrane. Identification of a fourth chain, alpha<sub>4</sub>, of type

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on:

March 6, 2002, 07:05:09 ; search time 18.75 Seconds

(without alignments)

125.149 Million cell updates/sec

Title: US-09-543-371-10\_COPY\_181\_244

Perfect score: 353

Sequence: 1 EFRASPLECHGRGTCNYNS..... KAGELEKIIISRCQVCMKKRH 64

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

To : number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 90%  
Maximum Match 100%

Listing first 1000 summaries

Database : SwissProt\_39.\*

Pred NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	353	100.0	1670	C34_HUMAN
2	323	91.5	471	1 C34_BOVIN

## ALIGNMENTS

RESULT	1	CA2_HUMAN	STANDARD;	PRT; 1670 AA.
IL	R34_HUMAN			
AC	001955;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	COLLAGEN ALPHA 3(IV) CHAIN PRECURSOR (GOODPASTURE ANTIGEN).			
GN	COL4A3.			
OS	Homo sapiens (Human).			
OC	Bukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
NCBI_TAXID=9606;				
[1]	SEQUENCE FROM N.A.			
RP	TISSUE=Kidney;			
RP	MEDLINE=931364904; PubMed=8083701;			
RA	Marikova M., Leinonen A., Mochizuki T., Tryggvason K., Reeders S.T.,			
RA	"Complete primary structure of the human alpha 3(IV) collagen chain in human tissues.", J. Biol. Chem. 269:23013-23017(1994).			
RP	[2]			
RP	REVISTONS.			
RA	Leinonen A.;			
RL	Submitted (OCT-1998) to the EMBL/GenBank/DDJB databases.			

- (3) [3] SEQUENCE OF 1386-1670 FROM N.A., AND PARTIAL SEQUENCE.  
 RP RX MEDLINE=93015826; PubMed=400251;  
 RA Quinones S., Bernali D., Garcia-Sogo M., Elena S.F., Saus J.;  
 RT "Exon/intron structure of the human alpha 3(IV) gene encompassing the Goodpasture antigen (alpha 3(IV)NC1). Identification of a potentially antigenic region at the triple helix/NC1 domain junction.";  
 RT J. Biol. Chem. 267:19780-19784(1992).  
 RN [4] SEQUENCE OF 1453-1670 FROM N.A.  
 RP MEDLINE=9135350; PubMed=1882840;  
 RA Morrison K.E., Mariyama M., Yang-Feng T.L., Reeders S.T.;  
 RT "Sequence and localization of a partial cDNA encoding the human alpha 3 chain of type IV collagen.";  
 RL An. J. Hum. Genet. 49:545-554(1991).  
 RN [5] SEQUENCE OF 1331-1670 FROM N.A.  
 RC TISSUE=Kidney;  
 RX TURNER N., Mason P.J., Brown R., Fox M., Povey S., Rees A.,  
 RA Pusey C.D.;  
 RT "Molecular cloning of the human Goodpasture antigen demonstrates it to be the alpha 3 chain of type IV collagen.";  
 RL J. Clin. Invest. 89:592-601(1992).  
 RN [6] SEQUENCE OF 1644-1670 FROM N.A.  
 RP RC TISSUE=Kidney;  
 RA Ding J.;  
 RL Submitted (JAN-1993) to the EMBL/GenBank/DDJB databases.  
 RN [7] SEQUENCE OF 1439-1670, AND ALTERNATIVE SPlicing.  
 RP RX TISSUE=Kidney;  
 RN MEDLINE=92147878; PubMed=173749;  
 RA Feng L., Xia Y., Wilson C.B.;  
 RT "Alternative splicing of the NC1 domain of the human alpha 3(IV) collagen gene. Differential expression of mRNA transcripts that predict three protein variants with distinct carboxyl regions.";  
 RT J. Biol. Chem. 269:2342-2348(1994).  
 RN [8] SEQUENCE OF 1-29 FROM N.A.  
 RP RX MEDLINE=94124597; PubMed=8294492;  
 RA Momota R., Sugimoto N., Oohashi T., Kigasawa K., Yoshioka H.,  
 RA Ninomiya Y.;  
 RT "Two genes, COLA3 and COLA4 coding for the human alpha3(IV) and alpha3(IV) collagen chains are arranged head-to-head on chromosome 2936.";  
 RT FEBS Lett. 424:11-16(1998).  
 RL [9] SEQUENCE OF 1-29 FROM N.A.  
 RN RP ALTERNATIVE SPlicing.  
 RX MEDLINE=93280184; PubMed=8505332;  
 RA Bernal D., Quinones S., Saus J.;  
 RT "The human mRNA encoding the Goodpasture antigen is alternatively spliced.";  
 RL J. Biol. Chem. 268:12090-12094(1993).  
 RN [10] SEQUENCE OF 1-29 FROM N.A.  
 RP VARIANT PRO-1474.  
 RX MEDLINE=95078827; PubMed=7987301;  
 RA Lemmink H.H., Mochizuki T., van den Heuvel L.P.W.J., Schroeder C.H.,  
 RA Barriontos A., Monnens L.A.H., van Oost B.A., Brunner H.G.,  
 RA Reeders S.T., Smets H.J.M.;  
 RT "Mutations in the type IV collagen alpha 3 (COL4A3) gene in autosomal recessive Alport syndrome";  
 RT Hum. Mol. Genet. 3:1269-1273(1994).  
 RN [11] SEQUENCE OF 1-29 FROM N.A.  
 RP VARIANT AS, AND VARIANTS  
 RX MEDLINE=21064696; PubMed=11134255;  
 RA Heider L., Arrendell L., Forestier L., Cohen-Solal L., Mollet G.,  
 RA Gutierrez B., Stavrou C., Gubler M.C., Antignac C.;  
 RT "Structure of the human type IV collagen gene COL4A3 and mutations in autosomal Alport syndrome";  
 RT J. Am. Soc. Nephrol. 12:97-105(2001);  
 RL CC -1 FUNCTION: TYPE IV COLLAGEN IS THE MAJOR STRUCTURAL COMPONENT OF GLOMERULAR BASEMENT MEMBRANES (GBM), FORMING A 'CHICKEN-WIRE'

CC	MESHWORK TOGETHER WITH LAMININS, PROTEOGLYCANS AND ENACTIN/
-1-	SUBUNIT: THERE ARE SIX TYPE IV COLLAGEN ISOFORMS. ALPHA 1(IV)-ALPHA 6(IV), EACH OF WHICH CAN FORM A TRIPLE HELIX STRUCTURE WITH 2 OTHER CHAINS TO GENERATE TYPE IV COLLAGEN NETWORK.
CC	-1- SUBCELLULAR LOCATION: CELL SURFACE (POTENTIAL).
CC	-1- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS; 1 (SHOWN HERE), 2/V AND 3/15; ARE PRODUCED BY ALTERNATIVE SPLICING. THEY DIFFER IN THEIR C-TERMINAL NC1 DOMAINS.
CC	-1- TISSUE SPECIFICITY: ALPHA 3 AND ALPHA 4 TYPE IV COLLAGENS ARE COLOCALIZED AND PRESENT ONLY IN BASEMENT MEMBRANES OF KIDNEY, EYE, COCHLEA, LUNG AND BRAIN.
CC	-1- DOMAIN: ALPHA CHAINS OF TYPE IV COLLAGEN HAVE A NONCOLLAGENOUS DOMAIN (NC1) AT THEIR C-TERMINUS, FREQUENT INTERRUPTIONS OF THE G-X-Y REPEATS IN THE LONG CENTRAL TRIPLE-HELICAL DOMAIN (WHICH MAY CAUSE FLEXIBILITY IN THE TRIPLE HELIX), AND A SHORT N-TERMINAL TRIPLE-HELICAL 7S DOMAIN.
CC	-1- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPETIDE REPEATING UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC	-1- PTM: THE ALTERNATIVE SPLICED FORM V CONTAINS AN ADDITIONAL N-LINKED GLYCOSYLATION SITE.
CC	-1- PTM: TYPE IV COLLAGENS CONTAIN NUMEROUS CYSTEINE RESIDUES WHICH ARE INVOLVED IN INNER- AND INTRAMOLECULAR DISULFIDE BONDING. 12 OF THESE, LOCATED IN THE NC1 DOMAIN, ARE CONSERVED IN ALL KNOWN TYPE IV COLLAGENS.
CC	-1- DISEASE: ANTIBODIES AGAINST THE NC1 DOMAIN OF ALPHA3(IV) MEDIATE THE HUMAN AUTOIMMUNE DISEASE, GOODPASTURE SYNDROME, WHICH IS CHARACTERIZED BY HEMATURIA AND PULMONARY HEMORRHAGE.
CC	-1- DISEASE: DEFECTS IN COL4A3 ARE ASSOCIATED WITH THE TYPE I AUTOSOMAL RECESSIVE FORM OF ALPORT SYNDROME, AN HEREDITARY GLOMERULONEPHROPATHY CHARACTERIZED BY PROGRESSIVE RENAL FAILURE, HEMATURIA AND DEAFNESS. THE RECESSIVE FORM OCCURS EQUALLY BETWEEN MALES AND FEMALES.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
DR	EMBL: X80031; CA56335_1; -.
DR	EMBL: M92933; AAA21610_1; -.
DR	EMBL: S55790; AAB19637_1; -.
DR	EMBL: M8179; AA51556_1; -.
DR	EMBL: L08650; AA52014_1; -.
DR	EMBL: U02319; AA18942_1; -.
DR	EMBL: AB088495; BRA23064_1; -.
DR	EMBL: M120070; -.
DR	MIM: 203780; -.
DR	MIM: 233450; -.
DR	InterPro: IPR001442; C4.
DR	InterPro: IPR000087; Collagen.
DR	Pfam: PF01413; C4; 2.
DR	Pfam: PF01391; Collagen; 21.
DR	ProDom: P00393; C4; 2.
DR	SMART: SM0011; C4; 2.
KW	Extracellular matrix; Connective tissue; Repeat; Hydroxylation; Glycoprotein; Basement membrane; Collagen; Signal; Cell adhesion; Alternative splicing; Polymorphism; Phosphorylation; Disease mutation; Apoptosis syndrome.
FT	SIGNAL 1 28 POTENTIAL.
FT	CHAIN 29 1670 COLLAGEN ALPHA 3(IV) CHAIN.
FT	DOMAIN 29 42 7S DOMAIN.
FT	DOMAIN 43 1438 TRIPLE HELICAL REGION.
FT	DOMAIN 1439 1670 NONHELICAL REGION (NC1) (GOODPASTURE ANTIGEN) (BY SIMILARITY).
FT	EPITOPE (RECOGNIZED BY GOODPASTURE ANTIBODIES).
FT	SITE 27 1426 CLEavage (BY COLLAGENASE) (BY SIMILARITY).
FT	SITE 27 1427
RP	SEQUENCE FROM N.A.
RP	SEQUENCE OF 227-258.
RP	SEQUENCE OF 91093146; PubMed=1985905;
RA	Morrison K.E., Germino G.G., Reeder S.T.,
RT	"Use of the polymerase chain reaction to clone and sequence a cDNA encoding the bovine alpha 3 chain of type IV collagen.", J. Biol. Chem. 266:34-39(1991).
RL	[2]
RN	[1]
RP	SEQUENCE OF 90202779; PubMed=2318822;
RA	Gunwar S., Saus J., Noelen M.E., Hudson B.G.,
RT	"Glomerular basement membrane. Identification of a fourth chain, alpha 4, of type IV collagen.",

```

Query Match      91.5%; Score 323; DB 1; Length 471;
Best Local Similarity 92.1%; Pred. No. 4e-33;
Matches 58; Conservative 3; Mismatches 2; Indels 0; Gaps 0

```

THIS PAGE BLANK (USPTO)

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on:

March 6, 2002, 07:04:44 ; Search time 50.56 Seconds

(without alignments)

185.155 Million cell updates/sec

Title: US-09-543-371-10\_COPY\_181\_244

Perfect score: 353

Sequence: 1 EFRASPFLECHGRGTCNYYS..... KAGELEKIIISRCOVCMKKRH 64

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 14627329 residues

T number of hits satisfying chosen parameters: 7

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000Post-processing: Minimum Match 90%  
Maximum Match 100%  
Listing first 1000 summaries

Database : SPTREMBL\_17:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	353	100.0	212	6 Q28512 macaca mula
2	353	100.0	245	4 Q9NYC4 homo sapien
3	353	100.0	1670	4 Q9BQ72 homo sapien
4	348	98.6	230	11 Q63122 rattus norv
5	348	98.6	246	11 Q61435 mus musculu
6	348	98.6	1669	11 Q9qzs0 mus musculu
7	331	93.8	212	6 Q28567 ovis aries

## ALIGNMENTS

RESULT	1	PRELIMINARY;	PRT;	212 AA.
Q28512				
ID	Q28512			
AC	028512;			
DT	01-NOV-1996 (TREMBREL_01, Created)			
DT	01-NOV-1996 (TREMBREL_01, Last sequence update)			
DT	01-JUN-2001 (TREMBREL_17, Last annotation update)			

DE	ALPHA-3 TYPE IV COLLAGEN (FRAGMENT).
GN	COL4A3.
OS	Macaca mulatta (Rhesus macaque).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidiae;
OC	Cercopithecinae; Macaca.
OX	NBII_TaxID=9544;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE-RENAL CORTEX;
RA	Turner A.N., Ryan J.J., Derry C.J., Cashman S.J., Katbamna I.,
RA	Mason P.J., Pusey C.D.; Submitted (MAR 1996) to the EMBL/GenBank/DBJ databases.
RL	EMBL; 147280; AAAI1861.1; -.
DR	INTERPRO; IPR00304; RRM.
DR	INTERPRO; IPR001442; C4.
DR	PRODOM; PD003923; C4; 2.
DR	PROSITE; PS00030; RRM_RNP_1; UNKNOWN_1.
DR	SMART; SM00111; C4; 2.
NON_TER	NON_TER
FT	NON_TER
SQ	SEQUENCE
	212 AA; 23469 MW; 4BC574A64E357E64 CRC64;
	Query Match 100.0%; Score 353; DB 6; Length 212;
	Best Local Similarity 100.0%; Pred. No. 1.8e-36; 0; Mismatches 0; Indels 0; Gaps 0;
	Matches 64; Conservative 0;
QY	1 EFRASPFLECHGRGTCNYYSNSYFWLNLPERMFRKPIPSVKAGELEKIIISRCOVCM 60
Db	149 EFRASPFLECHGRGTCNYYSNSYFWLNLPERMFRKPIPSVKAGELEKIIISRCOVCM 208
QY	61 KKRH 64
Db	209 KKRH 212
RESULT	2
ID	Q9NYC4
AC	Q9NYC4
DT	01-OCT-2000 (TREMBREL_15, Created)
DT	01-OCT-2000 (TREMBREL_15, Last sequence update)
DE	TUMSTATIN (FRAGMENT).
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NBII_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Masshima Y., Colorado P.C., Torre A., Holthaus K.A., Grunkemeyer J.A., Erickson M.D., Hooper H., Xiao Y., Stillman I.E., Kalluri R.; RT "Distinct anti-tumor properties of a type IV collagen domain derived from basement membrane." J. Biol. Chem. 2000; 275(20):14000-14004.
RA	EMBL; AR25851; AAF72632.1; -.
DR	INTERPRO; IPR001442; C4.
DR	INTERPRO; IPR000504; RRM.
DR	PRODOM; PP0143; C4; 2.
DR	SMART; SM00111; C4; 2.
DR	PROSITE; PS00030; RRM_RNP_1; UNKNOWN_1.
FT	NON_TER
SQ	SEQUENCE
	245 AA; 26952 MW; 1EE5028354D9A57D CRC64;
	Query Match 100.0%; Score 353; DB 4; Length 245;
	Best Local Similarity 100.0%; Pred. No. 2.1e-36; 0; Mismatches 0; Indels 0; Gaps 0;
	Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 EFRASPFLECHGRGTCNYYSNSYFWLNLPERMFRKPIPSVKAGELEKIIISRCOVCM 60

RESULT 3

ID Q9BQT2 PRELIMINARY; PRT; 1670 AA.

AC Q9BQT2; DR EMBL; AJ288532; CAC36101.1; JOINED.

DT 01-JUN-2001 (TREMBLrel. 17, Created)  
01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DE ALPHA3 TYPE IV COLLAGEN.

GN COLA3?

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

RN [1]

RP SEQUENCE FROM N.A.  
MEDLINE=21064690; PubMed=11134255;

RA Heidet L., Arrendel C., Forrestier L., Cohen-Solal L., Mollet G., Gutierrez B., Stavrou C., Gubler M.C., Antignac C.,  
"Structure of the human type IV collagen gene COLA3 and mutations in the autosomal Alport syndrome.", J. Am. Soc. Nephrol. 12:97-106 (2001).

RT DR EMBL; AJ288487; CAC36101.1; -

DR EMBL; AJ288488; CAC36101.1; JOINED.

DR EMBL; AJ288489; CAC36101.1; JOINED.

DR EMBL; AJ288490; CAC36101.1; JOINED.

DR EMBL; AJ288491; CAC36101.1; JOINED.

DR EMBL; AJ288492; CAC36101.1; JOINED.

DR EMBL; AJ288493; CAC36101.1; JOINED.

DR EMBL; AJ288494; CAC36101.1; JOINED.

DR EMBL; AJ288495; CAC36101.1; JOINED.

DR EMBL; AJ288496; CAC36101.1; JOINED.

DR EMBL; AJ288497; CAC36101.1; JOINED.

DR EMBL; AJ288498; CAC36101.1; JOINED.

DR EMBL; AJ288499; CAC36101.1; JOINED.

DR EMBL; AJ288500; CAC36101.1; JOINED.

DR EMBL; AJ288501; CAC36101.1; JOINED.

DR EMBL; AJ288502; CAC36101.1; JOINED.

DR EMBL; AJ288503; CAC36101.1; JOINED.

DR EMBL; AJ288504; CAC36101.1; JOINED.

DR EMBL; AJ288505; CAC36101.1; JOINED.

DR EMBL; AJ288506; CAC36101.1; JOINED.

DR EMBL; AJ288507; CAC36101.1; JOINED.

DR EMBL; AJ288508; CAC36101.1; JOINED.

DR EMBL; AJ288509; CAC36101.1; JOINED.

DR EMBL; AJ288510; CAC36101.1; JOINED.

DR EMBL; AJ288511; CAC36101.1; JOINED.

DR EMBL; AJ288512; CAC36101.1; JOINED.

DR EMBL; AJ288513; CAC36101.1; JOINED.

DR EMBL; AJ288514; CAC36101.1; JOINED.

DR EMBL; AJ288515; CAC36101.1; JOINED.

DR EMBL; AJ288516; CAC36101.1; JOINED.

DR EMBL; AJ288517; CAC36101.1; JOINED.

DR EMBL; AJ288518; CAC36101.1; JOINED.

DR EMBL; AJ288519; CAC36101.1; JOINED.

DR EMBL; AJ288520; CAC36101.1; JOINED.

DR EMBL; AJ288521; CAC36101.1; JOINED.

DR EMBL; AJ288522; CAC36101.1; JOINED.

DR EMBL; AJ288523; CAC36101.1; JOINED.

DR EMBL; AJ288524; CAC36101.1; JOINED.

DR EMBL; AJ288525; CAC36101.1; JOINED.

DR EMBL; AJ288526; CAC36101.1; JOINED.

DR EMBL; AJ288527; CAC36101.1; JOINED.

DR EMBL; AJ288528; CAC36101.1; JOINED.

DR EMBL; AJ288529; CAC36101.1; JOINED.

DR EMBL; AJ288530; CAC36101.1; JOINED.

RN [1]

RQ RESULT 4

ID 063122 PRELIMINARY; PRT; 230 AA.

AC 063122; DR EMBL; AJ288533; CAC36101.1; JOINED.

DT 01-NOV-1995 (TREMBLrel. 01, Created)  
01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DE ALPHA-3 TYPE IV COLLAGEN (FRAGMENT).

GN COLA3?

OS Rattus norvegicus (Rat).

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

RN [1]

RP SEQUENCE FROM N.A.  
STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY CORTEX;  
RA Turner N., Ryan J.J., Derry C.J., Cashman S.J., Katbama I.,  
Mason P.J., Pusey C.D.; DR EMBL; L47781; AAB72338.2; -  
RL Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.  
DR InterPro; IPR000504; RRM.  
DR InterPro; IPR001442; C4.  
DR Pfam; PF0413; C4; 2.  
DR ProDom; PD00323; C4; 2.  
DR PROSITE; PS0030; RRM\_RNP\_1; UNKNOWN\_1.  
DR SMART; SM0011; C4; 2.  
FT NON\_TER 1  
FT NON\_TER 230 230 AA; 230 MW; 29549E25314CC056 CRC64;  
SQ SEQUENCE 230 AA; 25398 MW;

Query Match 98 %; Score 348; DB 11; Length 230;  
Best Local Similarity 96.9%; Pred. No. 8.5e-36;  
Matches 62; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 | EFRASPFLECHGRGTCNYNSYSFWLASLNPERMFRKPISTVKAGELEKIIISRCQVM 60  
Db 167 | EFRASPFLECHGRGTCNYNSYSFWLASLNPERMFRKPISTVKAGELEKIIISRCQVM 226

QY 61 KKRH 64

Db 242 KKRH 245

RQ RESULT 5

ID Q61435 PRELIMINARY; PRT; 246 AA.

AC Q61435; DR EMBL; AJ288537; CAC36101.1; JOINED.

RQ RESULT 6

ID Q61435 PRELIMINARY; PRT; 246 AA.

AC Q61435; DR EMBL; AJ288538; CAC36101.1; JOINED.

DT 01-NOV-1996 (TREMBLREL. 01, Created)  
 DT 01-NOV-1996 (TREMBLREL. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMBLREL. 17, Last annotation update)  
 DE PROCOLLAGEN TYPE IV, ALPHA 3 (COLLAGEN IV ALPHA 3 CHAIN) (FRAGMENT).  
 GN COL4A3.  
 OS MUS musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TAXID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/C;  
 RX MEDLINE=95050957; PubMed=7962065;  
 RA Miner J.H.; Sanes J.R.;  
 RT "Collagen IV alpha 3, alpha 4, and alpha 5 chains in rodent basal  
 laminae: sequence, distribution, association with laminins, and  
 development switches";  
 RL J. Cell Biol. 121:879-891(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RP STRAIN=BALB/C;  
 RA Miner J.H.;  
 RL Submitted (JUL-1994) to the EMBL/GenBank/DBBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/C;  
 RA Miner J.H.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBBJ databases.  
 DR EMBL; 235166; CAR84529; 1; -. MGD; MGJ:104688; Col4a3.  
 DR InterPro; IPR00142; C4.  
 DR InterPro; IPR000504; RRM.  
 DR Pfam; PRO01413; C4; 2.  
 DR ProDom; PD003923; C4; 2.  
 DR SMART; SM00111; C4; 2.  
 DR PROSITE; PS00030; RRM\_RNP\_1; UNKNOWN\_1.  
 FT NON\_TER 1  
 SQ SEQUENCE 246 AA; 26993 MW; A9B5434F5836F324 CRC64;

Query Match 98.6%; Score 348; DB 11; Length 1669;  
 Best Local Similarity 96.9%; Pred. No. 9.1e-36; Length 246;  
 Matches 62; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EFRASPFLECHGRGTCNYISNSYSFWLASLNPERMFRRKPISTVKAGDELEKITSRCQVCM 60  
 Db 1606 EFRASPFLECHGRGTCNYISNSYSFWLASLNPERMFRRKPISTVKAGDELEKITSRCQVCM 1665  
 QY 61 KKRH 64  
 Db 1666 KKRH 1669

RESULT 7  
 ID Q28567 PRELIMINARY PRT; 212 AA.  
 AC Q28567;  
 DT 01-NOV-1996 (TREMBLREL. 01, Created)  
 DT 01-NOV-1996 (TREMBLREL. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMBLREL. 17, Last annotation update)  
 DE COL4A3.  
 GN ALPHA-3 TYPE IV COLLAGEN (FRAGMENT).  
 OS Ovis aries (Sheep).  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=ENERAL CORTEX;  
 RA Turner A.N.; Ryan J.J.; Derry C.J.; Cashman S.J.; Katbanna I.;  
 RA Mason P.J.; Pusey C.D.;  
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBBJ databases.  
 DR EMBL; L4782; AA091904; 1; -. MGD; MGJ:104688; Col4a3.  
 DR InterPro; IPR000504; RRM.  
 DR InterPro; IPR00142; C4.  
 DR Pfam; PRO01413; C4; 2.  
 DR ProDom; PD003923; C4; 2.  
 DR PROSITE; PS00030; RRM\_RNP\_1; UNKNOWN\_1.  
 FT SMART; SM00111; C4; 2.  
 FT NON\_TER 1  
 SQ SEQUENCE 212 AA; 23417 MW; 0F5839FCB81BDD8C CRC64;

Query Match 93.8%; Score 331; DB 6; Length 212;  
 Best Local Similarity 92.1%; Pred. No. 1.1e-33; Length 212;  
 Matches 58; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 EFRASPFLECHGRGTCNYISNSYSFWLASLNPERMFRRKPISTVKAGDELEKITSRCQVCM 60  
 Db 149 EYERANPFLECHGRGTCNYISNSYSFWLASLNPERMFRRKPISTVKAGDELEKITSRCQVCM 208  
 QY 61 KKR 63  
 Db 209 KRR 211

RT mouse model of alport syndrome.;"  
 RL Genomics 61:113-124 (1999); DR Genomics 61:113-124 (1999);  
 EMBL; AP169387; AAD50491; -. MGD; MGJ:104688; Col4a3.  
 DR InterPro; IPR00142; C4.  
 DR InterPro; IPR00087; Collagen.  
 DR InterPro; IPR000404; RRM.  
 DR Pfam; PRO01413; C4; 2.  
 DR Pfam; PRO01391; Collagen; 21.  
 DR ProDom; PD003923; C4; 2.  
 DR SMART; SM00111; C4; 2.  
 DR PROSITE; PS00030; RRM\_RNP\_1; UNKNOWN\_1.  
 DR Collagen. KW Sequence. 1669 AA; 161769 MW; 30976E59739A47B2 CRC64;

Query Match 98.6%; Score 348; DB 11; Length 1669;  
 Best Local Similarity 96.9%; Pred. No. 6.9e-35; Length 212;  
 Matches 62; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EFRASPFLECHGRGTCNYISNSYSFWLASLNPERMFRRKPISTVKAGDELEKITSRCQVCM 60  
 Db 1606 EFRASPFLECHGRGTCNYISNSYSFWLASLNPERMFRRKPISTVKAGDELEKITSRCQVCM 1665  
 QY 61 KKRH 64  
 Db 1666 KKRH 1669

Wed Mar 6 07:07:42 2002

us-09-543-371-10\_copy\_181\_244.rspt

Page 4

Search completed: March 6, 2002, 07:04:44  
Job time: 590 sec



**KW** Goodpasture syndrome; type IV collagen; alpha3 chain; human.  
**XX**  
**OS** Homo sapiens.  
**XX**  
**PN** US6007980-A.  
**XX** 28-DEC-1999.  
**PD**  
**XX** 07-OCT-1998; 98US-0167364.  
**XX** 30-NOV-1990; 90US-0621091.  
**PR** 07-MAR-1995; 95US-039889.  
**XX** (UNIV ) UNIV KANSAS MEDICAL CENT.  
**PA** (UYA ) UNIV YALE.  
**XX**  
**PI** Hudson BG, Reeder ST, Morrison KE;  
**XX** DR WPI; 2000-096371/08.  
**N-PSDB;** AAZ46729.

**PT** Diagnosing and treating Goodpasture syndrome using a peptide derived from type IV collagen - Disclosure; Columns 23-26; 26pp; English.

**CC** The invention provides a method of detecting Goodpasture antibodies in the fluid of a patient by contacting it with a peptide comprising at most 218 amino acids of the human alpha3 chain type IV collagen that contains the fragment shown in AAV56785. The methods are useful for the diagnosis and treatment of Goodpasture syndrome. The present sequence represents the carboxy terminal noncollagenous domain of the human alpha3 chain of type IV collagen.

**CC** Sequence 218 AA;

**SQ**

Query Match	100.0%	Score	353;	DB	21;	Length	218;
Best Local Similarity	100.0%	Pred. No.	9e-39				
Matches	64;	Conservative	0;	Mismatches	0;	Indels	0;
QY	1	EFRAFPFLCHGRGTCNYYSNSYNSFWLASSINPERMFRRPKIPSTVKAGELEKILISRCQVCM	60	QY	1	EFRAFPFLCHGRGTCNYYSNSYNSFWLASSINPERMFRRPKIPSTVKAGELEKILISRCQVCM	60
Db	155	efraspflechgrgtnyyssnsysfwlasinpermfrkripstvakelekiisrcqvcm	214	Db	182	efraspflechgrgtnyyssnsysfwlasinpermfrkripstvakelekiisrcqvcm	241
QY	61	KKRH	64	QY	61	KKRH	64
	215	kkrh	218				

**RESULT** 3

**ID** AAY67942 standard; Protein; 245 AA.

**AC** AAY67942;

**XX**

**DT** 03-APR-2000 (first entry)

**XX** Human type IV collagen alpha 3 chain protein sequence SEQ ID NO:10.

**KW** Human; type IV collagen; anti-angiogenic; angiogenesis; cancer; benign tumour; rheumatoid arthritis; diabetic retinopathy; psoriasis; ocular angiogenesis disease; Osler-Webber Syndrome; telangiectasia; myocardial angiogenesis; plaque neovascularisation; angiofibroma; atherosclerosis; scleroderma; hypertrophic scar; cat scratch disease; contraception; obesity.

**XX** Homo sapiens.

**OS** WO965940-A1.

**PN** PD 23-DEC-1999.

**XX**

**PF** 17-JUN-1999; 99WO-US13737.

**XX** 17-JUN-1998; 98US-0089689.

**PR** 25-MAR-1999; 99US-0126175.

**XX** (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

**PA**

**XX** DR WPI; 2000-097708/08.

**N-PSDB;** AAZ257158.

**PT** Anti-angiogenic proteins comprising the NC1 domain of the alpha 1, 2 or 3 chain of type IV collagen used in, e.g. treatment of benign tumors and rheumatoid arthritis -

**XX** PS Claim 32; Fig 16B; 117pp; English.

**CC** The present sequence represents the human type IV collagen alpha 3 chain. The present invention describes an isolated protein chosen from the NC1 domain of the alpha 1, alpha 2 or alpha 3 chains of type IV collagen or a fragment, analogue, derivative or mutant, which has anti-angiogenic properties. The anti-angiogenic proteins, multimers and chimeras are useful for inhibiting angiogenic activity in mammalian tissue, especially for treating diseases chosen from angiogenesis-dependent cancers, benign tumours, rheumatoid arthritis, diabetic retinopathy, psoriasis, ocular angiogenesis diseases, Osler-Webber Syndrome, myocardial angiogenesis, plaque neovascularisation, telangiectasia, haemophilic joints, angiofibroma, wound granulation, intestinal adhesions, atherosclerosis, scleroderma, hypertrrophic scars, cat scratch disease, Helicobacter pylori ulcers, dialysis graft vascular access stenosis, contraception and obesity. The compositions can be used to inhibit a disease characterised by angiogenic activity, in conjunction with radiation therapy, chemotherapy or immunotherapy.

**CC** Sequence 245 AA;

**SQ**

Query Match	100.0%	Score	353;	DB	21;	Length	245;
Best Local Similarity	100.0%	Pred. No.	1e-38;				
Matches	64;	Conservative	0;	Mismatches	0;	Indels	0;
QY	1	EFRAFPFLCHGRGTCNYYSNSYNSFWLASSINPERMFRRPKIPSTVKAGELEKILISRCQVCM	60	QY	1	EFRAFPFLCHGRGTCNYYSNSYNSFWLASSINPERMFRRPKIPSTVKAGELEKILISRCQVCM	60
Db	182	efraspflechgrgtnyyssnsysfwlasinpermfrkripstvakelekiisrcqvcm	241	Db	182	efraspflechgrgtnyyssnsysfwlasinpermfrkripstvakelekiisrcqvcm	241
QY	61	KKRH	64	QY	61	KKRH	64
	242	kkrh	245				

**RESULT** 4

**ID** AAY31993 standard; Protein; 268 AA.

**AC** AAY31993;

**XX**

**DT** 05-JAN-2000 (first entry)

**XX** Type IV collagen NC1 domain alpha-3 monomer.

**DE**

**XX** Type IV collagen; NC1 domain; non-collagenous domain; human; angiogenesis; tumour; metastasis; therapy; diabetic retinopathy; rheumatoid arthritis; retinal neovascularization; choroidal neovascularization; macular degeneration; corneal neovascularization; retinopathy of prematurity; corneal graft rejection; neovascular glaucoma; retrolental fibroplasia; epidemic keratoconjunctivitis; vitamin A deficiency; contact lens overwear; atopic keratitis; superior limbic keratitis; pterygium keratitis sicca; sores; acne rosacea; phylectenulosis; syphilis; Mycobacteria infection; lipid degeneration; chemical burn; ulcer; herpes simplex infection; Herpes zoster infection; protozoan infection; Kaposi's sarcoma;

CC	herpes zoster infections; protozoan infections; Kaposi's sarcoma;
CC	marginal keratolysis; trauma; systemic lupus; polyarteritis;
KW	Mooren ulcer; Terrien's marginal degeneration; marginal keratolysis;
KW	Wegener's sarcoidosis; scleritis; Steven's Johnson disease;
KW	radial keratotomy; sickle cell anaemia; sarcoid;
KW	pseudoxanthoma elasticum; Paget's disease; vein occlusion;
KW	artery occlusion; carotid obstructive disease; chronic uveitis;
KW	chronic vitritis; Lyme's disease; Eales disease; Bechets disease;
KW	myopia; optic pit; Stargart' disease; pars planitis;
KW	chronic retinal detachment; hypersensitivity syndrome; toxoplasmosis;
KW	post-laser complication; fibrovascular tissue proliferation;
KW	haemangioma; Osler-Weber-Rendu; AIDS; ocular neovascular disease;
KW	osteoarthritis; chronic inflammation; Crohn's disease;
KW	ulcerative colitis; psoriasis; atherosclerosis; pemphigoid.
XX	OS
OS	Homo sapiens.
XX	Synthetic.
PF	Location/qualifiers
FT	Key peptide
FT	1..17 "Bm40 signal peptide"
FT	Protein
FT	18..268 "/note= "mature protein"
FT	Peptide
FT	18..25 "/note= "affinity tag"
FT	26..268 "/note= "NCL alpha-3 monomer"
PR	W09949885-A2.
XX	PN
PD	07-OCT-1999.
PF	XX
PR	26-MAR-1999; 99WO-US06445.
PR	27-MAR-1998; 98US-0079783.
PR	29-OCT-1998; 98US-0106170.
XX	PA (UNIV ) UNIV KANSAS MEDICAL CENT.
PI	Hudson BG, Sarras MP;
XX	DR WPI; 1999-601297/51.
XX	N-PSDB; AAZ20091.
PT	Inhibition of angiogenesis with non-collagenous alpha chain monomer
PT	useful for treating e.g. tumor growth or metastasis,
PT	neovascularisation, etc.
XX	PS Disclosure: Fig 17c; 56pp; English.
CC	X This sequence represents a recombinant type IV collagen
CC	non-collagenous (NCL) domain alpha-3 polypeptide composed of a
CC	Bm40 signal sequence (which is cleaved from the mature protein) to
CC	facilitate protein secretion, and a mature protein comprising an
CC	affinity tag (facilitates purification and identification of the
CC	material) and the alpha-1 chain monomer. The invention provides
CC	methods and kits for inhibiting angiogenesis, tumour growth and
CC	metastasis, and endothelial cell interaction with the extracellular
CC	matrix, each method comprising contacting the tumour or animal
CC	tissue with 1 or more isolated type IV collagen NCL alpha chain
CC	monomer(s) selected from the group consisting of alpha-1, alpha-2,
CC	alpha-3 and alpha-6 NCL chain monomers (see Aay31991-96). The
CC	monomers can be produced via recombinant protein expression. The
CC	polynucleotides and polypeptides are used to treat an angiogenesis-
CC	mediated disorder or condition, especially selected from solid and
CC	retinal neovascularization, diabetic retinopathy, rheumatoid arthritis,
CC	retinal degeneration, corneal neovascularization, macular
CC	degeneration, corneal graft rejection, neovascular glaucoma, retrobulbar
CC	fibroplasia, epidemic keratoconjunctivitis, vitamin A deficiency,
CC	contact lens overwear, atopic keratitis, superior limbic keratitis,
CC	peringual keratitis, sicca, sognres, acne rosacea, phylectenulosis,
CC	syphilis, mycobacteria infections, lipid degeneration, chemical
CC	burns, bacterial ulcers, fungal ulcers, herpes simplex infections,
XX	OS
XX	Sequence 268 AA.
XX	RESULT 5
TD	Query Match 100.0%; Score 353; DB 20; Length 268;
XX	Best Local Similarity 100.0%; Pred. No. 1..1e-38;
XX	Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
AC	AY97155;
XX	DR 12-FEB-2001 (first entry)
XX	DE Human alpha3(IV)NCL protein sequence.
XX	XX Type IV collagen alpha chain monomer; human; inhibitor; angiogenesis;
XX	KW tumour growth; integrin receptor; carcinoma; sarcoma; rhabdomyosarcoma;
XX	KW retinoblastoma; Ewing Sarcoma; neuroblastoma; osteosarcoma; leukaemia;
XX	KW diabetic retinopathy; rheumatoid arthritis; neovascularisation;
XX	KW muscular degeneration; corneal graft rejection; vitamin A deficiency;
XX	KW atopic keratitis; Mycobacteria infection; chemical burn; sarcoid;
XX	KW kaposi's sarcoma; sickle cell anaemia; carotid obstructive disease;
XX	KW chronic inflammation; psoriasis; therapy; alpha3(IV)NCL.
OS	OS Homo sapiens.
XX	W020059532-A1.
PN	XX
XX	PD 12-OCT-2000.
PF	31-MAR-2000; 2000WO-US08678.
XX	PR 01-APR-1999; 99US-0127391.
PR	XX
XX	(BIOS-) BIOSTRATUM INC.
PA	XX
PI	Brooks P, Hudson B;
XX	WPI; 2000-664962/64.
DR	N-PSDB; AAA80993.
XX	XX
PR	Use of antagonists of specific integrin receptors for inhibiting
PR	angiogenesis, tumour growth or metastases, or endothelial cell
XX	interactions with the extracellular matrix -
XX	PS Disclosure: Fig 17c; 78pp; English.

CC This sequence is a human type IV collagen alpha chain monomer, designated alpha(IV)NC1. The invention relates to a method for inhibiting angiogenesis, tumour growth or metastases, or endothelial cell interactions with the extracellular matrix, comprising contacting the cells or tissue with a polypeptide composition containing antagonists of specific integrin receptors. The methods and the antagonists are useful for inhibiting angiogenesis, tumour growth or metastases, or endothelial cell interaction with the extracellular matrix. The antagonists are also useful for treating diseases and conditions with accompanying undesired angiogenesis, e.g. solid and blood borne tumours (e.g. melanomas, carcinomas, sarcomas, rhabdomyosarcoma, retinoblastoma, Ewing's sarcoma, neuroblastoma, osteosarcoma or leukaemia). These are also applicable to treating non-tumourigenic diseases and conditions with accompanying undesired angiogenesis, e.g. diabetic retinopathy, rheumatoid arthritis, retinal neovascularisation, choroidal neovascularisation, muscular degeneration, corneal graft rejection, vitamin A deficiency, atopic keratitis, Mycobacteria infections, chemical burns, Kaposi's sarcoma, sickle cell anaemia, sarcoid, carotid obstructive disease, post-laser complications, chronic inflammation or psoriasis.

Sequence 268 AA;

Query Match

Best Local Similarity 100.0%; Score 353; DB 21; Length 268;

Matches 64; Conservativeness 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EFRASPFLLEGHRGICNYNSYSWLASINPERMFRKIPSTVKAGELEKTIISRCQVCM 60

Db 205 efraspflechrgtgcnyvnsyswlas[inpermfrkipstvkagelekiisrcqvc]m 264

Qy 61 KKRH 64

Db 265 kkrh 268

RESULT 6

AAR79164 standard; Protein: 218 AA.

AC AAR79164;

XX DT 22-DEC-1995 (first entry)

DE Partial sequence of human alpha 3 chain of type IV collagen.

KW Type IV collagen; alpha 3 chain; Alport syndrome; COL4A3 gene.

Homo sapiens.  
US5424408-A.

PN XX

PD 13-JUN-1995.

XX PF 30-NOV-1990; 900S-0621091.

PR 30-NOV-1990; 900S-0621091.

XX PR 30-NOV-1990; 900S-0621091.

PA (UNIV ) UNIV KANSAS MEDICAL CENT.

PA (UYYA ) UNIV YALE.

PI Hudson BG, Morrison KE, Reenders ST;

DR WPT; 1995-262631/34.

DR N-PSDB; AAQ96291.

PT cDNAs encoding human or bovine alpha-3 type 4 collagen peptide(s) - useful for detection and therapeutic removal of antibodies associated with Goodpasture syndrome

PT disclosure; columns 7-10; 33pp; English.

CC Using the PCR with primers derived from each end of the known 27 AA residue bovine alpha 3 (IV) collagen protein sequence, a 68 bp bovine genomic fragment was amplified. This fragment was then used to a bovine lens cDNA library and a 1.5 kb partial cDNA clone was obtd. (clone KMC15). This encodes 238 residues of the triple helical

CC residue bovine alpha 3 (IV) collagen protein sequence, a 68 bp bovine genomic fragment was amplified. This fragment was then used to a bovine lens cDNA library and a 1.5 kb partial cDNA clone was obtd. (clone KMC15). This encodes 238 residues of the triple helical CC collagenous domain and all 233 residues of the C-terminal non- CC collagenous (NC1) domain of the alpha 3 (IV) chain. This bovine CC cDNA clone was used to screen a human kidney cDNA library and a CC 2.7 kb human cDNA clone (clone KMC21) was obtd. This clone encodes CC 218 residues of the NC1 domain and a portion of the 3' UTR region CC of the human alpha 3 (IV) chain. The COL4A3 gene localises to CC chromosome 2 and therefore mutations in COL4A3 cannot be CC responsible for Alport syndrome which is X-linked. An isolated CC and substantially pure nt. having the sequence in AAQ96291 is claimed.

Sequence 218 AA;

Query Match 95.8%; Score 338; DB 16; Length 218;

Best Local Similarity 96.9%; Pred. No. 8.8e-37; Matches 62; Conservativeness 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EFRASPFLLEGHRGICNYNSYSWLASINPERMFRKIPSTVKAGELEKTIISRCQVCM 60

Db 155 efpaspflechrgtgcnyvnsyswlas[inpermrsrkipstvkagelekiisrcqvc]m 214

Qy 61 KKRH 64

Db 215 kkrh 218

RESULT 7

AAR79163 standard; Protein: 471 AA.

ID AAR79163

XX AC AAR79163;

XX DT 22-DEC-1995 (first entry)

DE Partial sequence of bovine alpha 3 chain of type IV collagen.

KW Type IV collagen; alpha 3 chain.

OS Bos taurus.

PN US5424408-A.

XX PD 13-JUN-1995.

XX PF 30-NOV-1990; 900S-0621091.

XX PR 30-NOV-1990; 900S-0621091.

XX PA (UNIV ) UNIV KANSAS MEDICAL CENT.

PA (UYYA ) UNIV YALE.

XX PI Hudson BG, Morrison KE, Reenders ST;

DR WPT; 1995-262631/34.

DR N-PSDB; AAQ96290.

PT cDNAs encoding human or bovine alpha-3 type 4 collagen peptide(s) - useful for detection and therapeutic removal of antibodies associated with Goodpasture syndrome

PT disclosure; Columns 5-8; 33pp; English.

CC Using the PCR with primers derived from each end of the known 27 AA residue bovine alpha 3 (IV) collagen protein sequence, a 68 bp bovine genomic fragment was amplified. This fragment was then used to a bovine lens cDNA library and a 1.5 kb partial cDNA clone was obtd. (clone KMC15). This encodes 238 residues of the triple helical



THIS PAGE BLANK (USPTO)

Gencore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on:

March 6, 2002, 06:55:22 ; Search time 26.9 Seconds

(without alignments)

53.539 Million cell updates/sec

Title:

US-09-543-371-10-COPY\_181\_244

Perfect score:

353 EFRASPFLECHGRGTCNYYS..... KAGELEKIIISRCQVCMKKRH 64

Sequence:

BLOSUM62

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched:

212252 seqs, 22503292 residues

number of hits satisfying chosen parameters:

6

Minimum DB seq length:

0

Maximum DB seq length:

200000000

Post-processing:

Minimum Match 90%

Maximum Match 100%

Listing first 1000 summaries

Database :

Issued Patents AA:\*

```

1: /con2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_5/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /con2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_5/ptodata/2/1aa/PCUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

US-09-167-364-25

Result No.	Score	Query Length	DB ID	Description
1	353	100.0	218	2 US-09-399-889-25
2	353	100.0	218	3 US-09-167-364-25
3	353	100.0	218	4 US-09-439-897-4
4	323	91.5	471	2 US-09-399-889-24
5	323	91.5	471	3 US-09-167-364-24
6	323	91.5	471	4 US-09-439-889-2

RESULT 2  
US-09-167-364-25  
; Sequence 25, Application US/09167364  
; Patient No. 600780  
; GENERAL INFORMATION:  
; APPLICANT: Reeder, Stephen T  
; APPLICANT: Morrison, Karen E  
; TITLE OF INVENTION: Alpha-3 Chain Type IV Collagen Polypeptides  
; FILE REFERENCE: 951263B  
; CURRENT APPLICATION NUMBER: US-09-167-364  
; CURRENT FILING DATE: 1998-10-07  
; EARLIER APPLICATION NUMBER: 08/399889  
; EARLIER FILING DATE: 1995-03-07  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO: 25  
; LENGTH: 218  
; TYPE: PRT  
; ORGANISM: Human  
; US-09-167-364-25

RESULT 3  
US-09-439-897-4  
; Sequence 4, Application US/09439897  
; Patient No. 6277558  
; GENERAL INFORMATION:  
; APPLICANT: Hudson, Billy G  
; TITLE OF INVENTION: Alpha-3 Chain Type IV Collagen Polypeptides  
; FILE REFERENCE: 95-1263-C  
; CURRENT APPLICATION NUMBER: US-09/439,897  
; CURRENT FILING DATE: 1999-11-12  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO: 4  
; LENGTH: 218  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-439-897-4

RESULT 1  
US-09-399-889-25  
; Sequence 25, Application US/08399889B  
; Patient No. 5973120  
; GENERAL INFORMATION:  
; APPLICANT: Reeder, Stephen T  
; APPLICANT: Morrison, Karen E  
; APPLICANT: Hudson, Billy G  
; TITLE OF INVENTION: Alpha-3 Chain Type IV Collagen Polypeptides  
; FILE REFERENCE: 951263A  
; CURRENT APPLICATION NUMBER: US-08/399,889B  
; CURRENT FILING DATE: 1995-03-07  
; EARLIER APPLICATION NUMBER: 07/621091  
; EARLIER FILING DATE: 1990-11-30  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: Patentin Ver. 2.0

Query Match 100.0%; Score 353; DB 4; Length 218;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-39; Mismatches 0;  
 Matches 64; Conservative 0; Indels 0; Gaps 0;

QY 1 EFRASPFLECHGRGTCNYYSNSYSFWLASLNPERMFRKPKIPSTVKAGELEKITSRCQVCM 60  
 Db 155 EFRASPFLECHGRGTCNYYSNSYSFWLASLNPERMFRKPKIPSTVKAGELEKITSRCQVCM 214

QY 61 KKRH 64  
 Db 215 KRRH 218

RESULT 4

US-08-399-889-24

Sequence 24, Application US/08399889B  
 Patent No. 597120

GENERAL INFORMATION:

APPLICANT: Reeders, Stephen T

APPLICANT: Morrison, Karen E

APPLICANT: Hudson, Billy G

TITLE OF INVENTION: Alpha 3 Chain Type IV Collagen Polypeptides

FILE REFERENCE: 951263A

CURRENT APPLICATION NUMBER: US/08/399, 889B

CURRENT FILING DATE: 1995-03-07

EARLIER APPLICATION NUMBER: 07/621091

EARLIER FILING DATE: 1990-11-30

NUMBER OF SEQ ID NOS: 25

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 24

LENGTH: 471

TYPE: PRT

ORGANISM: Calf

US-08-399-889-24

Query Match 91.5%; Score 323; DB 4; Length 471;  
 Best Local Similarity 92.1%; Pred. No. 4.1e-35; Mismatches 2;  
 Matches 58; Conservative 3; Indels 0; Gaps 0;

QY 1 EFRASPFLECHGRGTCNYYSNSYSFWLASLNPERMFRKPKIPSTVKAGELEKITSRCQVCM 60  
 Db 408 EFRASPFLECHGRGTCNYYSNSYSFWLASLDPKMRKPKIPSTVKAGELENTISRCQVCM 467

QY 61 KKR 63  
 Db 468 KMR 470

RESULT 6

US-09-439-897-2

Sequence 2, Application US/09439897  
 Patent No. 6277558

GENERAL INFORMATION:

APPLICANT: Hudson, Billy G

TITLE OF INVENTION: Alpha-3 Chain Type IV Collagen Polypeptides

FILE REFERENCE: 95-1263-C

CURRENT APPLICATION NUMBER: US/09/439, 897

CURRENT FILING DATE: 1999-11-12

NUMBER OF SEQ ID NOS: 65

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2

LENGTH: 471

TYPE: PRT

ORGANISM: Bos taurus

US-09-439-897-2

Query Match 91.5%; Score 323; DB 4; Length 471;  
 Best Local Similarity 92.1%; Pred. No. 4.1e-35; Mismatches 2;  
 Matches 58; Conservative 3; Indels 0; Gaps 0;

US-09-167-364-24

Sequence 24, Application US/09167364  
 Patent No. 6007980

GENERAL INFORMATION:

APPLICANT: Reeders, Stephen T

APPLICANT: Morrison, Karen E

APPLICANT: Hudson, Billy G

TITLE OF INVENTION: Alpha-3 Chain Type IV Collagen Polypeptides

FILE REFERENCE: 951263B

CURRENT APPLICATION NUMBER: US/09/167, 364

CURRENT FILING DATE: 1998-10-07

EARLIER APPLICATION NUMBER: 08/399889

EARLIER FILING DATE: 1995-03-07

NUMBER OF SEQ ID NOS: 25

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 24

LENGTH: 471

TYPE: PRT

ORGANISM: Calf

US-09-167-364-24

Search completed: March 6, 2002, 06:55:22

Job time: 73 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

A;Residues: 1331-1524, 'I', 1526-1670 <TUR>  
A;Cross-references: GB;M81379  
A;Experimental source: kidney

## OM protein - protein search, using sw model

Run on: March 6, 2002; 06:54:09 : Search time 31.16 Seconds

Indicates /sec

Perfect score: 1052

URANGQUELLIGSCLORETTI.....KAGELEKISKUÖVMKKH 1911

Scoring table: BLOSUM62

THE JOURNAL OF CLIMATE

The number of hits satisfying chosen parameters: 3

Millennium DB seq length: 0

Maximum Match 100%

listing first 1000 summaries

Database : PIR\_68\_\*\*

```
2:     pir2: *
```

PAPERS OF THE AMERICAN ACADEMY IN BERLIN

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

600

NO.	Score	Match	Length	DB	ID	Description
1	1052	100.0	1670	1	CGHU3B	collagen alpha 3(I)
2	993	94.4	246	1	I48302	collagen alpha 3(I)
3	988	93.9	471	2	A39024	collagen alpha 3(I)

## ALIGNMENTS

**RESULT**  
**CGH03B**  
 collagen alpha 3(IV) chain precursor, long splice form - human  
 N; Alternate names: Goodpasture antigen; procollagen alpha 3(IV) chain long splice form  
 C; Species: Homo sapiens (man)  
 C; Date: 28-Oct-1994 #sequence\_revision 03-oct-1995 #text\_change 22-Jun-1999  
 C; Accession: A54763; A43928; A44043; A45971; A39786  
 R; Miyazawa, M.; Leironen, A.; Mochizuki, T.; Tryggvason, K.; Reeders, S. T.  
*J. Biol. Chem.*, 269, 23013-23017, 1994  
 A; Title: Complete primary structure of the human alpha3(IV) collagen chain. Coexpression  
 A; Reference number: A54763; MUID:94364994  
 A; Accession: A54763  
 A; Molecule type: mRNA  
 A; Residues: 1-1670 <MAR>  
 A; Cross-references: GB-x0031; NID:9577563; PID:9577564  
 A; Experimental source: kidney  
 R; Turner, N.; Mason, P.J.; Brown, R.; Fox, M.; Povey, S.; Rees, A.; Pusey, C.D.  
*J. Clin. Invest.*, 89, 592-601, 1992  
 A; Title: Molecular cloning of the human Goodpasture antigen demonstrates it to be the a  
 A; Reference number: A43928; MUID:9214878  
 A; Accession: A43928

Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORAHQDGLTGSLCQRFNTMPFLFCNYNDVCASRNDYSWLTSPALMPNMAPITGR 60  
Db 1480 ORAHQDGLTGSLCQRFNTMPFLFCNYNDVCASRNDYSWLTSPALMPNMAPITGR 1539

QY 61 ALEPYISRCTVCEGPAIATAVHSQTTDIPCPHGWISLWKGSFIMTSAGSEGTCGALA 120  
Db 1540 ALEPYISRCTVCEGPAIATAVHSQTTDIPCPHGWISLWKGSFIMTSAGSEGTCGALA 1599

QY 121 SPGSCLEEFRAFPFLECHGRTGTCNYNSYSWFLASLNPERMFRKIPSTVKAGELEKII 180  
Db 1600 SPGSCLEEFRAFPFLECHGRTGTCNYNSYSWFLASLNPERMFRKIPSTVKAGELEKII 1659

QY 181 SRCQVCMKKR 191  
Db 1660 SRCQVCMKKR 1670

LT 2  
I:18302  
collagen alpha 3(IV) chain - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Accession: I:18302; S47278  
R;Miner, J.H.; Saies, J.R.  
J;Cell Biol. 127, 879-891, 1994  
A;Title: Collagen IV alpha 3, alpha 4, and alpha 5 chains in rodent basal laminae: sequence  
A;Reference number: A54979; MUID:95050957  
A;Accession: I:18302  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-246 <RESTR>  
A;Cross-references: EMBL:235166; NID:9535197; PID:9535198  
C;Superfamily: collagen alpha 1(IV) chain

Query Match 94.4%; Score 993; DB 2; length 246;  
Best Local Similarity 92.7%; Pred. No. 4.6e-88; Indels 0; Gaps 0;  
Matches 177; Conservative 10; Mismatches 4; Indels 0; Gaps 0;

QY 1 ORAHQDGLTGSLCQRFNTMPFLFCNYNDVCASRNDYSWLTSPALMPNMAPITGR 60  
Db 56 KRAHGQDGLTGSLCQRFNTMPFLFCNYNDVCASRNDYSWLTSPALMPNMAPITGR 115

QY 61 ALEPYISRCTVCEGPAIATAVHSQTTDIPCPHGWISLWKGSFIMTSAGSEGTCGALA 120  
Db 116 ALEPYISRCTVCEGPAIATAVHSQTTDIPCPQDWISLWKGSFIMTSAGSEGTCGALA 175

QY 121 SPGSCLEEFRAFPFLECHGRTGTCNYNSYSWFLASLNPERMFRKIPSTVKAGELEKII 180  
Db 176 SPGSCLEEFRAFPFLECHGRTGTCNYNSYSWFLASLNPERMFRKIPSTVKAGELEKII 235

QY 181 SRCQVCMKKR 191  
Db 236 SRCQVCMKKR 246

RESULT 3  
A:9024  
collagen alpha 3(IV) chain - bovine (fragment)  
C;Species: Bos primigenius taurus (cattle)  
C;Accession: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 13-Aug-1999  
C;Date: 04-Dec-1992  
C;Accession: A:9024; S20672; S17802; A:35167; C39419; S13747; S20815  
R;Morison, K.E.; Germino, G.G.; Reeder, S.T.  
J;Biol. Chem. 266, 34-39, 1991  
A;Title: Use of the polymerase chain reaction to clone and sequence a cDNA encoding the  
A;Reference number: A:9024; MUID:9103146  
A;Accession: A:9024  
A;Molecule type: mRNA  
A;Residues: 1-471 <MOR>  
A;Cross-references: EMBL:63139; NID:9162886; PID:AAA62708.1; PID:9162887  
R;Butkowski, R.J.; Langeveld, J.P.M.; Wieslander, J.; Hamilton, J.; Hudson, B.G.

J. Biol. Chem. 262, 7874-7877, 1987  
A;Title: Localization of the Goodpasture epitope to a novel chain of basement membran  
A;Reference number: S18432; MUID:87222419  
A;Accession: S20672  
A;Molecule type: protein  
A;Residues: 227-228; X, 230-244 <BUR>  
R;Saus, J.; Wieslander, J.; Langenfeld, J.P.M.; Quiñones, S.; Hudson, B.G.  
J. Biol. Chem. 263, 13374-13380, 1988  
A;Title: Identification of the Goodpasture antigen as the alpha-3(IV) chain of collag  
A;Reference number: S17802; MUID:88330844  
A;Accession: S17802  
A;Molecule type: protein  
A;Residues: 227-228, X, 230-252, Y, 254 <SAU>  
R;Gunwar, S.; Saus, J.; Noeiken, M.E.; Hudson, B.G.  
J. Biol. Chem. 265, 5466-5469, 1990  
A;Title: Glomerular basement membrane. Identification of a fourth chain, alpha4, of t  
A;Reference number: A35167; MUID:90202779  
A;Accession: A35167  
A;Molecule type: protein  
A;Residues: 236-258 <CON>  
R;Gunwar, S.; Ballister, F.; Kalluri, R.; Timoneda, J.; Chonko, A.M.; Edwards, S.J.;  
R;Gunwar, S.; Ballister, F.; Kalluri, R.; Timoneda, J.; Chonko, A.M.; Edwards, S.J.;  
J;Biol. Chem. 266, 15318-15324, 1991  
A;Title: Glomerular basement membrane. Identification of dimeric subunits of the nonc  
A;Reference number: A39419; MUID:91322055  
A;Accession: C39419  
A;Molecule type: protein  
A;Residues: 236-255 <GU2>  
C;Superfamily: collagen alpha 1(IV) chain  
C;Keywords: basement membrane; cell binding; coiled coil; disulfide bond; duplication  
C;239-472/Domain: carboxyl-terminal nonhelical NC1 #status predicted <NC1>  
F;239-353/Domain: repeat NC1 #status predicted <NC1>  
F;232, 239/Domain: repeat NC1 #status predicted <NC1>  
F;232, 239/Modified site: hydroxyproline (Pro) #status experimental  
F;306-312, 417-423/Disulfide bonds: #status predicted  
F;306-312, 417-423/Disulfide bonds: #status predicted

Query Match 93.9%; Score 988; DB 2; Length 471;  
Best Local Similarity 93.2%; Pred. No. 2.7e-87; Indels 0; Gaps 0;  
Matches 177; Conservative 9; Mismatches 4; Indels 0; Gaps 0;

QY 1 ORAHQDGLTGSLCQRFNTMPFLFCNYNDVCASRNDYSWLTSPALMPNMAPITGR 60  
Db 281 EQAHQDGLTGSLCQRFNTMPFLFCNYNDVCASRNDYSWLTSPALMPNMAPITGR 340

QY 61 ALEPYISRCTVCEGPAIATAVHSQTTDIPCPHGWISLWKGSFIMTSAGSEGTCGALA 120  
Db 341 ALEPYISRCTVCEGPAIATAVHSQTTDIPCPAGWISLWKGSFIMTSAGSEGTCGALA 400

QY 121 SPGSCLEEFRAFPFLECHGRTGTCNYNSYSWFLASLNPERMFRKIPSTVKAGELEKII 180  
Db 401 SPGSCLEEFRAFPFLECHGRTGTCNYNSYSWFLASLNPERMFRKIPSTVKAGELEKII 460

QY 181 SRCQVCMKKR 190  
Db 461 SRCQVCMKKR 470

Search completed: March 6, 2002, 06:54:49  
Job time: 40 sec



CC MESHWORK TOGETHER WITH LAMININS, PROTEOGLYCANS AND ENACTIN/ NIDGEN. CC SUBUNITS: THERE ARE SIX TYPE IV COLLAGEN ISOFORMS, ALPHA 1(IV)- ALPHA 6(IV), EACH OF WHICH CAN FORM A TRIPLE HELIX STRUCTURE WITH 2 OTHER CHAINS TO GENERATE TYPE IV COLLAGEN NETWORK.

CC -!- SUBCELLULAR LOCATION: CELL SURFACE (POTENTIAL). FT SITE 791 793 FT SITE 996 998 FT SITE 1156 1156 FT SITE 1306 1308 FT SITE 1347 1347 FT SITE 1434 1434 FT CAREHYD 253 253 FT MOD,RES 1435 1435 FT MOD,RES 1437 1437 FT DISULFID 1460 1551 FT DISULFID 1493 1548 FT DISULFID 1505 1511 FT DISULFID 1570 1665 FT DISULFID 1604 1662 FT DISULFID 1616 1622 FT VARSPLIC 1586 1670

CC COCHLEA, LONG AND BRAIN. CC DOMAIN (NCI) AT THEIR C-TERMINUS, FREQUENT INTERRUPTIONS OF THE G-X-Y REPEATS IN THE LONG CENTRAL 'TRIPLE-HELICAL DOMAIN' (WHICH MAY CAUSE FLEXIBILITY IN THE TRIPLE HELIX), AND A SHORT N-TERMINAL TRIPLE-HELICAL 7S DOMAIN.

CC -!- PROLINE: TYPE IV COLLAGEN HAVE A NONCOLLAGENOUS UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.

CC -!- PTM: THE ALTERNATIVE SPLICED FORM V CONTAINS AN ADDITIONAL N-LINKED GLYCOSYLATION SITE.

CC -!- PTM: TYPE IV COLLAGENS CONTAIN NUMEROUS CYSTEINE RESIDUES WHICH ARE INVOLVED IN INTER- AND INTRAMOLECULAR DISULFIDE BONDING. 12 OF THESE, LOCATED IN THE NCI DOMAIN, ARE CONSERVED IN ALL KNOWN TYPE IV COLLAGENS.

CC -!- DISEASE: ANTIBODIES AGAINST THE NCI DOMAIN OF ALPHA3(IV) MEDIATE THE HUMAN AUTOIMMUNE DISEASE, GOODPASTURE SYNDROME, WHICH IS CHARACTERIZED BY HEMATURIA AND PULMONARY HEMORRHAGE.

CC -!- DISEASE: DEFECTS IN COL4A3 ARE ASSOCIATED WITH THE TYPE I AUTOSOMAL RECESSIVE FORM OF ALPORT SYNDROME, AN HEREDITARY HEMATURIA AND DEAFNESS. THE RECESSIVE FORM OCCURS EQUALLY BETWEEN MALES AND FEMALES.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR X80031; CAM5335; 1; . DR EMBL; M22993; AA21610; 1; . DR EMBL; M55790; AAB19637; 1; . DR EMBL; M81379; AA51556; 1; . DR EMBL; U08650; AA52044; 1; . DR EMBL; U02519; AA18942; 1; . DR EMBL; U02520; AA18943; 1; . DR EMBL; AB008495; BAA25064; 1; . DR MIM; 120070; . DR MIM; 233450; . DR InterPro; IPR01442; C4. DR InterPro; IPR000087; Collagen. DR Pfam; PF01413; C4; 2. DR Pfam; PF01391; Collagen; 21. DR Prodom; PD005923; C4; 2. DR SMART; SM00111; C4; 2. DR SMART; SM00111; C4; 2. DR Extracellular matrix; Connective tissue; Repeat; Hydroxylatation; KW Glycoprotein; Basement membrane; Collagen; Signal; Cell adhesion; KW Alternative splicing; Polymorphism; Phosphorylation; Disease mutation; KW Alport syndrome. FT SIGNAL 1 28 FT CHAIN 29 1670 FT DOMAIN 29 42 FT DOMAIN 43 1438 FT DOMAIN 1439 1670 FT DOMAIN 1427 1444 FT SITE 1426 1427

RESULT 2

	CA34_BOVIN	STANDARD:	PRT:	471 AA.
TD	CA34_BOVIN			
AC	028084;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	COLLAGEN ALPHA 3(IV) CHAIN (FRAGMENT).			
GN	COL4A3.			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Buteraria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;			
OC	Bovidae; Bovinae; Bos.			
OX	NCBIL_TaxID=9913;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Lens;			
RX	SBQQUE=Lens;			
RA	Medline=91093146; PubMed=1985905;			
RA	Morrison K.E., Germino G.G., Reeder S.T.;			
RT	"use of the polymerase chain reaction to clone and sequence a cDNA encoding the bovine alpha 3 chain of type IV collagen";			

FT SITE 1426 1427

RL J. Biol. Chem. 266:34-39(1991).  
 RN [2]  
 RP SEQUENCE OF 227-258.  
 RC TISSUE-Kidney;  
 RA Gunwar S., Saus J., Noeiken M.E., Hudson B.G.;  
 RT "Glomerular basement membrane; Identification of a fourth chain,  
 alpha 4, of type IV collagen";  
 RL J. Biol. Chem. 265:5466-5469(1990).  
 RN [3]  
 RP SEQUENCE OF 227-254.  
 RX Saus J., Wieslander J., Langeveld J.P.M., Quinones S., Hudson B.G.;  
 RA "Identification of the Goodpasture antigen as the alpha 3(IV) chain  
 of collagen IV";  
 RT MEDLINE=88330844; PubMed=2417661;  
 RL [4]  
 RP SEQUENCE OF 227-244.  
 RX MEDLINE=87222419; PubMed=2438203;  
 RA Butkowski R.J., Langeveld J.P.M., Wieslander J., Hamilton J.,  
 Hudson B.G.;  
 RT Localization of the Goodpasture epitope to a novel chain of basement  
 membrane collagen;  
 RL J. Biol. Chem. 262:7874-7877(1987).  
 CC -1- FUNCTION: TYPE IV COLLAGEN IS THE MAJOR STRUCTURAL COMPONENT OF  
 GLOMERULAR BASEMENT MEMBRANE (GBM), FORMING A 'CHICKEN-WIRE'  
 MESHWORK TOGETHER WITH LAMININS, PROTEOGLYCANS AND ENTACTIN/  
 NIDROGEN.  
 CC -1- SUBUNIT: THERE ARE SIX TYPE IV COLLAGEN ISOFORMS, ALPHA 1(IV)-  
 ALPHA 6(IV), EACH OF WHICH CAN FORM A TRIPLE HELIX STRUCTURE  
 WITH 2 OTHER CHAINS TO GENERATE TYPE IV COLLAGEN NETWORK.  
 CC -1- SUBCELLULAR LOCATION: CELL SURFACE (POTENTIAL).  
 CC -1- DOMAIN: ALPHA CHAINS OF TYPE IV COLLAGEN HAVE A NONCOLLAGENOUS  
 DOMAIN (NC1) AT THEIR C-TERMINUS, FREQUENT INTERRUPTIONS OF THE  
 G-X-Y REPEATS IN THE LONG CENTRUS, TRIPLE-HELICAL DOMAIN (WHICH MAY  
 CAUSE FLEXIBILITY IN THE TRIPLE HELIX), AND A SHORT N-TERMINAL  
 TRIPLE HELICAL 7S DOMAIN.  
 CC -1- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPETIDE REPEATING  
 UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.  
 CC -1- PTM: TYPE IV COLLAGENS CONTAIN NUMEROUS CYSTEINE RESIDUES WHICH  
 ARE INVOLVED IN INTER- AND INTRAMOLECULAR DISULFIDE BONDING. 12 OF  
 THESE, LOCATED IN THE NC1 DOMAIN, ARE CONSERVED IN ALL KNOWN TYPE  
 IV COLLAGENS.  
 CC -1- SIMILARITY: TO OTHER TYPE IV COLLAGENS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/announce/or\\_send\\_an\\_email\\_to\\_licensee@isb-sib.ch](http://www.isb-sib.ch/announce/or_send_an_email_to_licensee@isb-sib.ch)).

EMBL; M63139; AA62708; 1;  
 DR InterPro; IPR01442; C4.  
 DR InterPro; IPR000087; Collagen.  
 PRIM; PF01413; C4; 2.  
 PRIM; PF01391; Collagen; 4.  
 PRODom; P003923; C4; 2.  
 SMART; SMM0011; C4; 2.  
 DR SMART; SMM0011; C4; 2.  
 KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;  
 Glycoprotein; Basement membrane; Collagen; Cell adhesion.  
 FT NON\_TER 1 1  
 FT DOMAIN <1 238  
 FT DOMAIN 239 471  
 FT SITE 106 108  
 FT MOD\_RES 232 232  
 FT MOD\_RES 238 238  
 FT DISULFID 261 352  
 FT DISULFID 294 349  
 FT DISULFID 306 312  
 FT DISULFID 371 466  
 FT DISULFID 405 463

	FT	DISULFID	417	423	BY SIMILARITY.
	CONFICT	253	253	S -> Y	(IN REF. 3).
SO	SEQUENCE	471 AA;	47585 MW;	C03B66F14E7008DE	CRC64;
<b>Query Match</b>					
Qy	1	ORAHGDIGTGGSCLQRPTMPFLCNVNFDVNFASRNDSYKWLSTPALMPNMAPITGR	281	ALEPYISRCTVGEPATAIAVHSQTDIIPCPAGWISLWKGFSFIMETSAGSEGAGAOLA	120
Db	61	EQAHGDIGTGGSCLQRPTMPFLCNVNFDVNFASRNDSYKWLSTPALMPNMAPITGR	340	ALEPYISRCTVGEPATAIAVHSQTDIIPCPAGWISLWKGFSFIMETSAGSEGAGAOLA	400
Qy	121	SPGSCLEBRFRASPFLLEGHGRGICNYNSYSFWLASINPERMFKPPIPSTVKAGELEII	460	SPGSCLEBRFRASPFLLEGHGRGICNYNSYSFWLASIDPKMRKPPIPSTVKAGELEII	180
Db	401	SPGSCLEBRFRASPFLLEGHGRGICNYNSYSFWLASIDPKMRKPPIPSTVKAGELEII	470	SPGSCLEBRFRASPFLLEGHGRGICNYNSYSFWLASIDPKMRKPPIPSTVKAGELEII	181
Qy	181	SRQCQVCKMR	470	SRQCQVCKMR	470
Db	461	SRQCQVCKMR	470	SRQCQVCKMR	470

Search completed: March 6, 2002, 07:05:08 Job time: 584 sec

THIS PAGE BLANK (USPTO)